
Automatisierung der 3D-Rekonstruktion aus mehreren Tiefenbildern durch Optimierung prädizierter Qualitätsgewinne

Vom Fachbereich Informatik
der Technischen Universität Darmstadt
genehmigte

DISSERTATION

zur Erlangung des akademischen Grades eines
Doktor-Ingenieurs (Dr.-Ing.)

von

Dipl.-Inform. Konrad Klein

geboren am 17. Februar 1972 in Würzburg

Referenten der Arbeit: Prof. Dr.-Ing. Georgios Sakas
Prof. Dr.-Ing. Matthias Teschner
Prof. Dr.-Ing. Michael Goesele
Tag der Einreichung: 29. April 2008
Tag der Prüfung: 16. Juni 2008

Darmstadt, 2008
D17

Vorwort

Die vorliegende Arbeit entstand während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Fraunhofer-Institut für Graphische Datenverarbeitung. Danken möchte ich an dieser Stelle allen, die zur Entstehung dieser Arbeit beigetragen haben.

Besonderer Dank gilt meinem Doktorvater Herrn Prof. Dr.-Ing. Georgios Sakas für die Überlassung des Themas, die Betreuung der Arbeit und nicht zuletzt auch für sein großes Engagement bei der Einwerbung der Fördermittel, durch welche die im Rahmen der Arbeit verwendeten Geräte beschafft werden konnten. Ebenso gilt mein Dank Herrn Prof. Dr.-Ing. Dr. h.c. mult. Dr. E.h. Hon. Prof. mult. J. L. Encarnação für die Unterstützung der Arbeit. Herrn Prof. Dr.-Ing. Matthias Teschner und Herrn Prof. Dr.-Ing. Michael Goesele danke ich für ihre Bereitschaft, die Gutachten zu übernehmen.

Ebenfalls danken möchte ich meinen ehemaligen Kollegen und Vorgesetzten am Fraunhofer-Institut für Graphische Datenverarbeitung sowie an der Gemeinsamen Forschungsstelle der Europäischen Kommission in Ispra, Italien, für die angenehmen Arbeits- und Forschungsbedingungen, die ich während meiner Zeit an beiden Instituten vorgefunden habe. Dr.-Ing. Peter J. Neugebauer danke ich für die beständige Diskussionsbereitschaft über Themen der 3D-Rekonstruktion. Auch den übrigen Kollegen, insbesondere Vítor Sequeira und Hagen Schumann, gilt mein Dank für die Unterstützung der Arbeit. Außerdem danke ich allen an der Arbeit beteiligten Diplomanden, studentischen Hilfskräften und Praktikanten für ihr Engagement.

Nicht zuletzt möchte ich meiner Frau und meinem Vater für kritisches Korrekturlesen sowie meiner Familie für das aufgebrachte Verständnis danken, wenn sie hinter der Arbeit an der Dissertation zurückstehen mußte. Danke!

Kurzfassung

Mit Methoden der optischen Meßtechnik können die Oberflächen realer Objekte abschnittsweise flächenhaft digitalisiert werden. Die Planung, Umsetzung und Integration der Teilaufnahmen zur vollständigen Vermessung von Objekten erfordert jedoch viel Zeit und Erfahrung. Daher bilden Verfahren zur automatischen Aufnahmeplanung unter Berücksichtigung der Datenqualität (vor allem im Sinne der Abtastdichte) den Gegenstand der vorliegenden Dissertation. Als Ergebnis ermöglichen die vorgeschlagenen Verfahren die vollautomatische Digitalisierung initial unbekannter Objekte mit Hilfe geeigneter Roboter.

Es werden Lösungen für die beiden zentralen Teilprobleme der Aufnahmeplanung vorgeschlagen: nämlich für die Repräsentation des erst teilweise vermessenen Objekts, sowie für die Ableitung einer geeigneten Zielsetzung für die Aufnahmeplanung. Letztere nutzt die vorliegende Objektrepräsentation zur Prädiktion des Qualitätsgewinns, der von einem zusätzlichen Aufnahmevergang zu erwarten ist. Dabei wird gezeigt, wie sich die intrinsischen Eigenschaften des Sensors unter Verwendung computergraphischer Verfahren zur Planung nutzen lassen, indem gerenderte Darstellungen der Objektrepräsentation analysiert werden.

Ergänzend wird eine Methodik zur Evaluation der Ergebnisse vorgeschlagen. Ergebnisse aus dem Bereich der Gebäudevermessung und der Objektdigitalisierung vervollständigen die Arbeit.

Abstract

Methods of optical metrology enable the surface digitisation of real world objects by piecewise area scanning. However, the planning, acquisition, and integration of multiple images necessary to completely measure the surface of objects requires substantial time and experience. This thesis consequently focusses on techniques of automatic view planning for the acquisition of images, aiming at the highest achievable quality of the measurements (namely in terms of sampling density). As a result, the digitisation of initially unknown objects can be performed fully automatic using suitable robots.

This thesis presents solutions for the two main problems of view planning: the representation of the partly known object geometry, and the derivation of suitable criteria for the view planning process. The latter is achieved by using the available object representation to predict the increase in quality, which can be expected from an additional acquisition step. The thesis demonstrates in which way the intrinsic properties of the sensor can be exploited, namely using methods of computer graphics to analyse rendered views of the object representation.

In addition, a methodology for the evaluation of the resulting 3d geometries is proposed. Exemplary results documenting the automated measurement of buildings as well as small object digitisation complete the thesis.

Inhaltsverzeichnis

1	Einleitung	11
1.1	Motivation	11
1.2	Stand der Technik	12
1.3	Zielsetzung und Abgrenzung der Arbeit	15
1.4	Kapitelübersicht	17
2	Optische Meßtechnik und Automatisierung	19
2.1	Einleitung	19
2.2	Stand der Technik	20
2.3	Optische 3D-Meßtechnik	21
2.3.1	Unifokale Systeme	22
2.3.2	Triangulations-basierte Systeme	28
2.3.3	Multifokale Systeme	46
2.3.4	Abstraktionsgrad und Schnittstelle	47
2.4	Automatisierungstechnik	48
2.4.1	Stationäre Roboter	49
2.4.2	Mobile Roboter	56
2.4.3	Abstraktionsgrad und Schnittstelle	58
2.5	Zusammenfassung eigener Ergebnisse	58
3	Objektrepräsentation	61
3.1	Einleitung	61
3.2	Stand der Technik	62
3.3	Definition der teilvermessenen Objektoberfläche	63
3.4	Inkrementelle Konstruktion der Objektoberfläche	68
3.4.1	Vermaschung überlappender Dreiecksnetze	71
3.4.2	Punktrendering	72
3.4.3	Grenzflächen des eingesehenen Volumens	75
3.4.4	Repräsentation qualitativer Maßzahlen	77
3.5	Initiale Repräsentation	82

3.5.1	Interne Rekonstruktion	83
3.5.2	Externe Rekonstruktion	83
3.6	3D-Rekonstruktion aus mehreren Einzelbildern	85
3.6.1	Redundanznutzung in mehrfach vermessenen Bereichen	85
3.6.2	Oberflächendefinition und Triangulierung	86
3.7	Zusammenfassung eigener Ergebnisse	87
4	Ziel und Umsetzung der Aufnahmeplanung	89
4.1	Einleitung	89
4.2	Stand der Technik	91
4.3	Formulierung des Problems	92
4.3.1	Zielsetzung	92
4.3.2	Aktions- und Suchraum	95
4.3.3	Nebenbedingungen	98
4.4	Definition der Zielfunktion	99
4.5	Effiziente Berechnung der Zielfunktion	104
4.5.1	Grundprinzip	106
4.5.2	Abbildung der Zielfunktion auf graphische Basisfunktionen	107
4.6	Globale Optimierung der Aufnahmeparameter	117
4.6.1	Anforderungen	117
4.6.2	Reihenfolge der Funktionsevaluationen	118
4.6.3	Abtasten des Parameterraums	119
4.7	Abbruchkriterium des Gesamtverfahrens	127
4.8	Zusammenfassung eigener Ergebnisse	128
5	Bewertung und Ergebnisse	131
5.1	Einleitung	131
5.2	Evaluationskriterien	131
5.2.1	Überblick eigener Vorschläge	131
5.2.2	Volumetrisches Kriterium	132
5.2.3	Abtastdichte	133
5.3	Ergebnisse und Diskussion	135
5.3.1	Interne Rekonstruktion	135
5.3.2	Externe Rekonstruktion	142
5.4	Zusammenfassung eigener Ergebnisse	145
6	Zusammenfassung und Ausblick	149
A	Formeln	153
A.1	Berechnung der Sensorauflösung	153
Literatur		155

KAPITEL 1

Einleitung

1.1 Motivation

In den letzten Jahren wurde eine Vielzahl von Verfahren vorgeschlagen, die die Rekonstruktion dreidimensionaler Objekte und Szenen aus Tiefenbildern ermöglichen. In der Praxis erfordern diese Verfahren jedoch noch einen großen Anteil interaktiver Tätigkeit durch erfahrenes Bedienpersonal. Insbesondere die Akquisition der benötigten Tiefenbilder ist ein zeitaufwändiger Prozess, der von den meisten Benutzern nicht effizient durchgeführt werden kann.

Eine Alternative zur bisher üblichen manuellen Vorgehensweise bildet die Automatisierung des Rekonstruktionsprozesses. Die dafür benötigten Roboter stehen bereits seit langem für alle relevanten Objektgrößen zur Verfügung. Die bisher vorgeschlagenen Algorithmen zur Aufnahmeplanung sind jedoch nicht hinreichend skalierbar, um alle praxisrelevanten Anwendungsfälle abzudecken. Die Aufnahmeplanung bildet also einen limitierenden Faktor für die schnelle und kostengünstige Erzeugung von 3D-Modellen für computergraphische Anwendungen, für rapid-prototyping-Verfahren, und für die sogenannte *as-built*-Analyse von Bauwerken und Industrieanlagen.

Es besteht eine auffällige Verwandtschaft zwischen den Teilproblemen, die bei der Aufnahmeplanung auftreten, und zahlreichen Lösungsansätzen der Computergraphik. Beispielhaft genannt seien hier an erster Stelle die Sichtbarkeitsbestimmung und die Schattierungsrechnung, aber auch Methoden der *radiosity* und der Abschattungsrechnung stehen in engem Bezug zur Aufnahmeplanung für marktgängige 3D-Meßsysteme.

Eine wesentliche Motivation der Arbeit besteht daher darin, die bereits etablierten Methoden der Computergraphik auch für die Aufnahmeplanung nutzbar zu machen.

Die Bedeutung der Computergraphik ist an dieser Stelle nicht nur algorithmischer, sondern auch rechentechnischer Natur durch die breite Verfügbarkeit spezialisierter Hardware in Gestalt von 3D-Graphik-Subsystemen, die zur Beschleunigung des rechenintensiven Prozesses der Aufnahmeplanung genutzt werden kann.

Eine effizient ablaufende automatische Aufnahmeplanung eröffnet eine Vielzahl von zusätzlichen Anwendungsfeldern der 3D-Digitalisierung, die bisher wegen des notwendigen zeitaufwändigen Einsatzes hochqualifizierten Bedienpersonals aus Kostengründen nicht praktikabel waren. Hier ist vor allem an die Geometrie-Digitalisierung für 3D-Produktkataloge zu denken. Weitere Anwendungsfelder ergeben sich im medizintechnischen Bereich, etwa bei der Digitalisierung von Abformungen in der Dentaltechnik oder in der Otoplastik. In der industriellen Qualitätskontrolle erfordert das zunehmende Auftreten kleiner Fertigungslose bis hin zur „Losgröße 1“ geeignete Digitalisierungslösungen, die den Kostenanteil der Qualitätssicherung von der Losgröße entkoppeln. Weitere relevante Anwendungsfelder sind die Dokumentation in der Archäologie sowie die systematische Erkundung nicht mehr zugänglicher Gebäudeteile nach Katastrophen oder zur Kampfmittelräumung durch autonome mobile Roboter.

1.2 Stand der Technik

Die bisher vorgeschlagenen Verfahren zur Aufnahmeplanung für initial unbekannte 3D-Oberflächen basieren, wie auch der in der vorliegenden Arbeit verfolgte Ansatz, auf zwei Komponenten: einer geeigneten Beschreibung des Objekts, in der die bereits bekannten Oberflächenbereiche zusammen mit der möglichen Lage noch nicht eingesehener Bereiche repräsentiert sind, sowie einer Zielfunktion mit entsprechendem Optimierungsverfahren. Ausführliche Übersichten zur Thematik wurden unter anderem von Tarabanis et al. [105] sowie weniger lang zurückliegend von Scott et al. [97] veröffentlicht, in denen auch ausführlich auf die Planung basierend auf bekannten Geometrien (CAD-Modelle) eingegangen wird. Eine Klassifikation vorgeschlagener Verfahren zur Aufnahmeplanung ist in Abbildung 1.1 dargestellt.

Die Arbeit von Conolly [18] ist der erste bekannte Versuch, das Problem der Aufnahmeplanung zu lösen. Der bereits 1985 vorgeschlagene Ansatz basiert auf einer einfach skalierbaren Diskretisierung sowohl der Objektbeschreibung mittels eines binären Octrees als auch des Suchraums mittels eines sphärischen Rasters. Die Zielfunktion wird durch einen ray-casting-Ansatz aus der Octree-Repräsentation berechnet; letzterer wird inkrementell angepaßt. Wegen der zum Veröffentlichungszeitpunkt aufgetretenen exorbitanten Rechenzeiten wird alternativ ein vereinfachter Algorithmus vorgeschlagen, welcher eventuell auftretende Eigenverdeckungen kurzerhand ignoriert. Da die Entwicklung der Computergraphik heute effizientere Verfahren als das

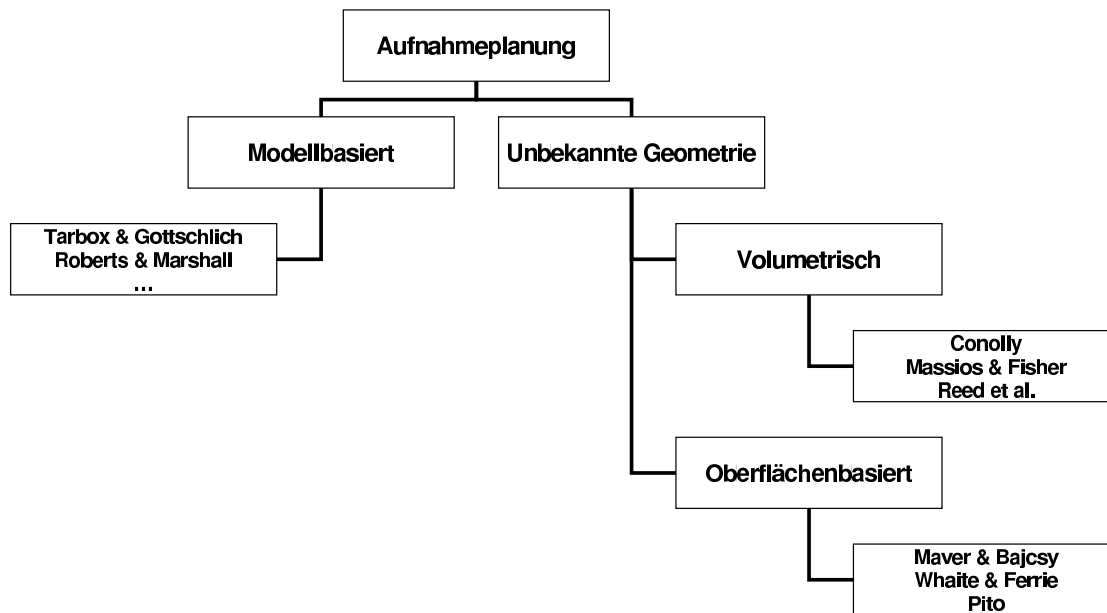


Abbildung 1.1: Klassifikation vorgeschlagener Verfahren (in Anlehnung an Scott et al. [97]).

binär-volumetrische ray casting ermöglicht, darf das Verfahren als überholt bezeichnet werden.

Der Ansatz von Whaite und Ferrie [114] bildet einen Sonderfall unter den Verfahren zur Aufnahmeplanung, da eine nicht-diskrete geometrische Beschreibung der Objektoberfläche die Grundlage des Verfahrens bildet. Die Autoren verwenden eine „Unsicherheits-Ummantelung“ („shell of uncertainty“), welche die tatsächliche (teilweise noch unbekannte) Oberfläche umschließt und durch sogenannte *superquadrics* repräsentiert wird. Die Bestimmung der nächsten optimalen Ansicht hat die Minimierung des Volumens dieser „Unsicherheits-Ummantelung“ zum Ziel. Der Ansatz ist nur für einfache 3D-Oberflächen geeignet, da sowohl die Berechnung der Objektrepräsentation als auch die Evaluierung potentieller Aufnahme-Parametersätze extrem rechenaufwändig sind.

Maver und Bajcsy [65] schlagen ein Verfahren vor, das auf einen triangulationsbasierten Laser-Lichtschnitt-Sensor spezialisiert ist und dessen Schattenbildung berücksichtigt. Das Verfahren optimiert zunächst die Orientierung des Triangulationsaufbaus, um aus der bereits verwendeten Richtung weitere Aufnahmen zu akquirieren, die verbleibende Lücken schließen. In einem zweiten Schritt werden dann weitere Aufnahmerichtungen bestimmt. Die Arbeit zeichnet sich durch eine detaillierte Unterscheidung der verschiedenen in einem triangulations-basierten Meßsystem auftretenden

Abschattungs-Situationen aus und verwendet einen Histogramm-Ansatz, um die nächste Aufnahme zu bestimmen.

Eine Variante des volumetrischen Ansatzes wird von Massios und Fisher [64] vorgeschlagen, bei der nicht nur die Vermessung bisher nicht eingesehener Oberflächenteile als Ziel verfolgt wird, sondern auch die erneute Vermessung bisher suboptimal vermessener Oberflächenteile. Es handelt sich also im weiteren Sinne um einen qualitätsbezogenen Ansatz.

Pito [79, 80] schlägt eine Ergänzung der bereits vermessenen Oberfläche durch sogenannte *void surfaces* vor, die den bisher vom Sensor eingesehenen Teil des Raumes begrenzen. Die Optimierung der Aufnahmeparameter erfolgt zweischrittig, indem zunächst ohne Verdeckungsrechnung diejenigen Teile des Suchraumes identifiziert werden, die mögliche Optima enthalten. Die Begriffe der *void surfaces* und der *void-Volumen* werden wegen ihrer Kürze und Übersetzungsresistenz auch in der vorliegenden Arbeit verwendet und in Abschnitt 3.3 generalisiert und detaillierter definiert. Der Kernpunkt des von Pito vorgeschlagenen Verfahrens ist die Repräsentation von Planungszwischenergebnissen im sogenannten "positional space", der mit jedem zusätzlichen Bild fortgeschrieben wird. Es erscheint allerdings fraglich, ob dieser Ansatz, der nur für einen niedrig aufgelösten zweidimensionalen Suchraum vorgeführt wird, auch für größere und höherdimensionale Suchräume geeignet ist.

Unter den Verfahren zur Aufnahmeplanung läßt sich eine eigene Klasse unterscheiden, welche die Oberflächen der durch Eigenverdeckungen nicht eingesehenen Volumina diskret analysiert. Reed et al. [87, 86] berechnen für jede dieser *void surfaces* ein Polyeder, der alle Punkte einschließt, von denen aus die jeweilige *void surface* sichtbar ist. Sequeira et al. [99] schlagen ein ähnliches Verfahren vor, das zusätzlich die Anzahl der benötigten Aufnahmepositionen minimiert. Während Reed et al. einen beliebigen Punkt innerhalb des Polyeders verwenden, schlagen Sequeira et al. eine Zielfunktion zur Optimierung der Aufnahmeparameter innerhalb des Polyeders vor. Roberts und Marshall [89] versuchen ebenfalls, mit einer minimalen Anzahl von Aufnahmevergängen auszukommen, indem sie eine Graphenstruktur in den Teilstücken der Oberfläche aufbauen, welche diejenigen Teilstücke durch eine Kante verbindet, die von einem gemeinsamen Punkt aus eingesehen werden können.

Die vollständige Evaluation des Suchraumes kann durch Verfahren dieser Klasse vermieden werden. Abhängig vom konkreten Anwendungsfall kann aber die Anzahl der zu berücksichtigenden *void surfaces* eine bedeutende Rolle spielen (einige der genannten Verfahren besitzen eine nicht-polynomielle Komplexität in der Anzahl der *void surfaces*), so daß die Verfahren auf geometrisch einfache Objekte beschränkt sind. Das Anwendungsgebiet liegt daher in der Planung auf Basis vorbekannter Geometrien durch CAD-Modelle im Rahmen von Qualitätssicherungsaufgaben. Die Problematik der nicht-polynomiellen Komplexität wird von Tarbox und Gottschlich [106] sowie von Scott et al. [96] thematisiert.

Eine weitere eigene Klasse wird durch Verfahren zur Rekonstruktion von Umgebungen unter freiem Himmel gebildet. Stellvertretend seien hier Blaer und Allen [8] sowie Teller [107] genannt.

1.3 Zielsetzung und Abgrenzung der Arbeit

In der Literatur werden bezüglich des Problems der Aufnahmeplanung in Abhängigkeit von der vorab bekannten geometrischen Information mehrere Varianten unterschieden. In der vorliegenden Arbeit wird der Fall einer anfangs unbekannten Geometrie mit vorgegebener Initialansicht thematisiert. Dabei werden die Parameter der folgenden Tiefenbildansichten iterativ innerhalb des folgenden, in Abbildung 1.2 dargestellten Zyklus' geplant:

1. Tiefenbildakquisition.
2. Integration des neu aufgenommenen Bildes in das rekonstruierte Modell.
3. **Aufnahmeplanung:** Bestimmung der Parameter der nächsten Akquisition oder Beendigung, wenn keine weitere Verbesserung des Rekonstruktionsergebnisses erreicht werden kann.
4. Bewegung des Akquisitionssystems.

Im Unterschied zum klassischen Ansatz der Aufnahmeplanung, der lediglich die Vollständigkeit der Vermessung zum Ziel hat, wird in der vorliegenden Arbeit ein Qualitätsmaß berücksichtigt. Das Ziel der Arbeit ist daher die automatische Vermessung von 3D-Oberflächen mit einer vorgegebenen Abtastdichte und einer vorgegebenen Genauigkeit des resultierenden Modells.

Der oben skizzierte iterative Ablauf eignet sich, die verschiedenen Aspekte der vorliegenden Arbeit sowie die betreffenden Zielsetzungen und Abgrenzungen zu strukturieren, die im folgenden beschrieben werden. Auf die nach der Erfassung der benötigten Tiefenbildaufnahmen folgende 3D-Rekonstruktion wird dagegen nicht weiter eingegangen, da hierzu im Rahmen der vorliegenden Arbeit keine eigenen Lösungsvorschläge beigetragen werden. Auf die Weiterverarbeitung der Tiefenbilder wird aber an zahlreichen Stellen Bezug genommen um die jeweilige Vorgehensweise bei der Planung zu motivieren und zu vergleichen.

Die Tiefenbildakquisition bildet einen für die Aufnahmeplanung wichtigen Aspekt der 3D-Rekonstruktion, da die Eigenschaften des Sensors bekannt sein müssen, um eine Ergebnis-Prädiktion im Rahmen der Planung durchführen zu können. Dabei ist eine gewisse Abstraktion vom konkret eingesetzten Gerät wünschenswert. In der vorliegenden Arbeit wird daher eine entsprechende Klassifizierung vorgeschlagen. Meßtechnische Details werden nur soweit beschrieben, wie sie beim Zusammenspiel des

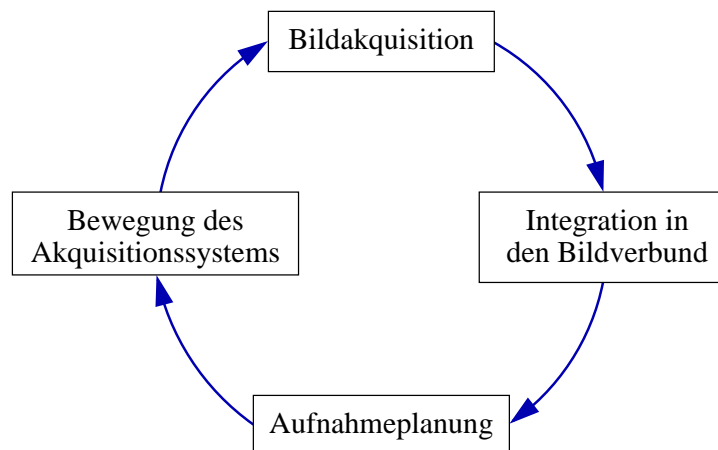


Abbildung 1.2: Die einzelnen, iterativ durchlaufenen Schritte des Aufnahmezyklus’.

Sensors mit dem Positionierungssystem sowie für die Prädiktion der zu planenden Aufnahme relevant sind.

Die Integration zusätzlicher Aufnahmen berührt einen der Kernpunkte der vorliegenden Arbeit. Die iterative Aufnahmeplanung erfordert eine effiziente, inkrementell ergänzbare Repräsentation der teilvermessenen Oberfläche. Das verwendete Kalkül ist die volumetrische Betrachtung der Sensorinformation, die neben den erfaßten Oberflächendaten auch die Aufnahmeposition und die Bildgeometrie auswertet. Hierzu werden neue Verfahren vorgeschlagen, und die zugrundeliegenden Überlegungen zur Repräsentation der Oberfläche und mit ihr assoziierter Maßzahlen werden im Detail beschrieben.

Die eigentliche Aufnahmeplanung basiert auf dieser im Mehrbildintegrationsschritt gewonnenen Objektrepräsentation und ist in zwei Teilaspekte unterteilt: zum einen wird ein Optimierungsverfahren auf Basis planungstheoretischer Strukturen mit einer geeigneten Zielfunktion formal vorgeschlagen, zum anderen bildet die effiziente Berechnung mit den dazu geeigneten Vereinfachungen einen Schwerpunkt der Arbeit. Die beabsichtigte Rückführung der hier notwendigen Rechenschritte auf computergraphische Basisfunktionen bildet einen weiteren Kernpunkt der vorliegenden Arbeit.

Die Bewegung des Akquisitionssystems mit Hilfe eines geeigneten Roboters wird in der vorliegenden Arbeit im Wesentlichen vorausgesetzt. Wichtige Teilprobleme wie zum Beispiel die Wegplanung mobiler Roboter gehen über den Rahmen der vorliegenden Arbeit hinaus. Zur Vereinfachung der Roboter-Anbindung wird aber eine Schnittstelle vorgeschlagen, die bereits vorhandene Verfahren zur Bewegungsplanung, wie sie in Industrierobotern typischerweise implementiert sind, sinnvoll integriert.

1.4 Kapitelübersicht

Im nachfolgenden zweiten Kapitel wird zunächst die eingesetzte Technik zur 3D-Vermessung und zu ihrer Automatisierung beschrieben und klassifiziert. Dabei bildet die Identifizierung geeigneter Abstraktionsgrade und entsprechender Schnittstellen den für die weiteren Kapitel wichtigsten Aspekt.

Die Kapitel 3 und 4 beinhalten die Kernaspekte der vorliegenden Arbeit. Im dritten Kapitel wird zunächst eine geeignete Repräsentation der anfänglich nur teilweise bekannten Objektgeometrie vorgeschlagen. Neben der Repräsentation der Oberfläche und der mit ihr assoziierten Maßzahlen bildet die inkrementelle Ergänzung dieser Repräsentation im Rahmen des iterativen Aufnahmeprozesses einen wichtigen Teil dieses Kapitels.

Im vierten Kapitel wird die eigentliche Planungsaufgabe beschrieben und ein Verfahren zu ihrer Lösung vorgeschlagen. Auch hier werden zunächst formal das Ziel und der Algorithmus konkretisiert, bevor für verschiedene Sensortypen die Abbildung rechenintensiver Teilschritte auf computergraphische Basisfunktionen hergeleitet wird. Damit wird neben der Möglichkeit, die Rechengeschwindigkeit durch Nutzung des Graphik-Subsystems zu erhöhen, vor allem durch die Reduzierung der Rechengenauigkeit im Rahmen des Übergangs auf die Bildauflösung der Graphikausgabe eine wesentliche Beschleunigung des Gesamtverfahrens erreicht. Ein weiterer wichtiger Aspekt ist die Optimierung der jeweiligen nächsten Aufnahmekonfiguration. Hierzu wird ein Ansatz vorgeschlagen, der in dem vorliegenden hochdimensionalen, aufgrund von diversen Nebenbedingungen aber nur schwer systematisch zu analysierenden Parameterraum eine bezüglich der verfügbaren Rechenzeit skalierbare Optimierung vornimmt.

Kapitel 5 thematisiert vor der Präsentation von Ergebnissen zunächst Kriterien und Analysemethoden, mit denen Verfahren der Aufnahmeplanung evaluiert werden können. Die vorgeschlagenen Verfahren sollen einen systematischen Vergleich mit anderen Aufnahmeplanungen ermöglichen. Zusammenfassung und Ausblick schließen in Kapitel 6 die Arbeit ab.

KAPITEL 2

Verfahren der optischen Meßtechnik und der Automatisierungstechnik

2.1 Einleitung

Die automatisierte 3D-Rekonstruktion setzt geeignete Geräte zur Datenakquisition voraus, die in diesem Kapitel beschrieben werden. Die Digitalisierung der Oberfläche soll dabei durch flächenhaft optisch messende Systeme erfolgen, deren relative Lage zum Objekt mit Hilfe einer Positioniereinrichtung vom Rechner gesteuert werden kann. Daraus ergibt sich für kleine Objekte als typischer Aufbau ein ortsfestes 3D-Meßsystem, vor dem das zu vermessende Objekt von einem Roboter bewegt wird. Für mittelgroße Objekte, deren Positionierung aufwändiger als die des Meßsystems ist, bleibt meist das Objekt ortsfest oder kann nur eingeschränkt (zum Beispiel auf einem Drehteller) positioniert werden, während das Meßsystem von einem Roboter bewegt wird. Im Architektur-Bereich und für sehr große Objekte ist im Regelfall ein mobiles Aufnahmesystem unumgänglich, welches ein Meßsystem mit entsprechender Reichweite bewegt.

Das in den folgenden Kapiteln beschriebene Verfahren zur Aufnahmeplanung und Automatisierung eignet sich prinzipiell für alle flächenhaft messenden Systeme; die wichtigsten Typen werden in Abschnitt 2.3 beschrieben. Eine für das entwickelte Verfahren wichtige Eigenschaft des Meßsystems ist der jeweilige Standort des Sensors bei der Erfassung eines Meßwertes. Daher bildet dieser Aspekt das Unterscheidungskriterium bei der Klassifizierung der unterschiedlichen Typen von Meßsystemen.

Während die Verfahren der optischen Meßtechnik im Rahmen der Aufnahmeplanung relativ detailliert berücksichtigt werden müssen, können die Verfahren zur Positionierung des Meßsystems in Abschnitt 2.4 vereinfacht beschrieben werden. Hier

kommt es im Wesentlichen auf die Berücksichtigung des Arbeitsbereichs an, während Details der Robotik wie Inverse Kinematik und Pfadplanung über den Rahmen der für die vorliegende Arbeit relevanten Themen hinausgehen und daher nur oberflächlich dargestellt werden.

2.2 Stand der Technik

Die optische Vermessung von sichtbaren Objektoberflächen ist bereits seit längerem Stand der Technik. Die im folgenden Abschnitt 2.3 gegebene Übersicht lehnt sich an die Arbeit von Besl aus dem Jahr 1989 an [7]; seitdem sind zwar zahlreiche Detailverbesserungen und Varianten der dort beschriebenen Verfahren publiziert, aber keine grundlegend neuen Meßverfahren vorgeschlagen worden [9, 15]. Besl klassifiziert Tiefenbildgebungs-Sensorik zum einen nach den eingesetzten Meßprinzipien, zum anderen nach der Geometrie des nativen Sichtfeldes und den erforderlichen Relativbewegungen zwischen Sensor und Objekt, die zur Erfassung eines zweidimensionalen Tiefenbildes beziehungsweise der kompletten Objektoberfläche notwendig sind. Auf diese Weise ergibt sich eine „natürliche Hierarchie“ [7] von Sensortypen:

1. Punkt-Sensoren, die zur Erzeugung einer bildartigen Meßpunktmenge eine in zwei Dimensionen parametrisierte Relativbewegung von Sensor und Objekt erfordern.
2. Zeilen-Sensoren, die eine eindimensional parametrisierte Relativbewegung zur Erzeugung einer bildartigen Meßpunktmenge erfordern.
3. Flächen-Sensoren, die direkt eine bildartige Meßpunktmenge liefern.
4. Mehrblickpunkt-Sensoren, die eventuell verbleibende Lücken eines Einzelbildes durch Hinzunahme weiterer Blickpunkte schließen.

Jeder Sensortyp innerhalb dieser Hierarchie kann durch Hinzunahme einer Ablenk- oder Positioniereinheit in ein System der nächsthöheren Hierarchiestufe überführt werden.

Während die ersten drei Hierarchiestufen intuitiv eine einfache Parametrisierungsmöglichkeit nahelegen, erscheint die von Besl vorgeschlagene Kategorie der Mehrblickpunkt-Sensoren als Bruch innerhalb der Hierarchie. Mathematisch ist die ursächliche Schwierigkeit der konsistenten Klassifizierung dadurch begründet, daß die bei der Oberflächenvermessung erfaßten Punktmengen in einer in den dreidimensionalen Raum eingebetteten, gekrümmten 2D-Oberfläche liegen. Daher können die Elemente der Punktmenge in null, einer oder zwei Dimensionen abschnittsweise geordnet vorliegen; eine höhere Dimensionalität ist nicht möglich.

Gleichzeitig bietet die abschnittsweise geordnete Repräsentation von Punktmengen gegenüber ungeordneten Punktmengen den Vorteil, daß die Oberfläche auf einfache Weise zwischen den Meßpunkten durch Interpolation stetig ergänzt werden kann. Innerhalb der von Besl vorgeschlagenen Hierarchie ergibt sich so ein Vorteil beim Übergang von Punkt- zu Zeilen-Sensoren sowie von Zeilen- zu Flächen-Sensoren, während der Übergang zu Mehrblickpunkt-Sensoren die Handhabung der entstehenden Punktmengen nicht vereinfacht, sondern zunächst erschwert, da redundant vermessene Bereiche geeignet berücksichtigt werden müssen.

Aus diesen Gründen wird in der vorliegenden Arbeit eine strikte Trennung vorgenommen, die zwischen dem System zur Oberflächenerfassung und dem Positioniersystem unterscheidet, obwohl das vorgeschlagene Gesamtsystem durchaus ein Mehrblickpunkt-Sensor im Besl'schen Sinne ist. In der vorliegenden Arbeit werden alle mechanischen Komponenten, die zur Erzeugung einer „bildartigen“, also zweidimensional parametrisierten Meßpunktmenge erforderlich sind, dem Meßsystem zugeordnet, die übrigen mechanischen Komponenten dagegen dem Positioniersystem.

2.3 Optische 3D-Meßtechnik

Mit Methoden der optischen Meßtechnik kann berührungslos geometrische Information über die einsehbaren Teile der Oberfläche eines vermessenen Objekts gewonnen werden. Diese geometrische Information liegt zunächst als Punktwolke vor, welche in vielen Fällen in einem zweidimensionalen Bildraster, einem sogenannten *Tiefenbild* organisiert ist.

Für die automatisierte Erfassung von Oberflächen mit Methoden der optischen Meßtechnik ist es naturgemäß notwendig, nicht nur die Geometrie der gemessenen Punkte, sondern auch die Standorte und Sichtlinien des Sensors relativ zur vermessenen Oberfläche zu berücksichtigen. Nur auf diese Weise kann die Sichtbarkeit einer zu vermessenden Oberfläche bei der Bewertung einer möglichen Aufnahmekonfiguration entschieden werden. Daher wird in der vorliegenden Arbeit die einfache Punktwolke als eine unvollständige Repräsentation des Meßergebnisses angesehen. Zusätzlich wird zu jedem Meßpunkt auch die Lage der genutzten Beobachtungspunkte benötigt. Neben der Oberflächengeometrie kann so auch eine volumetrische Information genutzt werden: zwischen dem vermessenen Oberflächenbereich und den Beobachtungspunkten kann kein verdeckendes Element lokalisiert sein; das betreffende Volumen muß also leer sein¹.

¹Hier und im Folgenden wird davon ausgegangen, daß sich keine transparenten Materialien (Glas, Wasser) im Meßbereich befinden.

Dieser bekanntermaßen leere Teil des Raumes wird mit dem Begriff des „eingesehenen Volumens“ bezeichnet und spielt eine entscheidende Rolle bei der Planung der weiteren Messungen. Zum einen kann nur für das eingesehene Volumen ausgeschlossen werden, daß ein bisher unzureichend oder noch gar nicht erfaßter Teil der Objektoberfläche von einem gegebenen Blickpunkt aus durch andere Objektteile verdeckt wird und daher nicht vermessen werden kann. Zum anderen kann das Meßsystem oftmals nur dann sicher an eine gewünschte Aufnahmeposition bewegt werden, wenn diese Bewegung im eingesehenen Volumen stattfindet, da auf diese Weise Kollisionen ausgeschlossen werden können.

Daher wird im Folgenden besonders auf die Lage und Anzahl der genutzten Beobachtungspunkte eingegangen. In Abhängigkeit vom Meßprinzip und vom Aufbau des Meßsystems treten unterschiedliche Anzahlen von Beobachtungspunkten bei der Erfassung eines Oberflächenpunktes beziehungsweise bei der Erfassung eines Tiefenbildes auf. Dies wird im Folgenden als Kriterium zur Klassifikation eingesetzt:

Unifokale Sensoren nutzen nur einen Beobachtungspunkt zur Erfassung eines Meßpunktes, der für alle Rasterpunkte eines Tiefenbildes identisch ist. Entsprechende Meßprinzipien werden in Abschnitt 2.3.1 beschrieben.

Triangulations-Sensoren nutzen mehrere (meist zwei) räumlich signifikant verschiedene Beobachtungspunkte, die für alle Rasterpunkte eines Tiefenbildes identisch sind. Sie sind Gegenstand von Abschnitt 2.3.2.

Multifokale Sensoren erfassen die einzelnen Meßpunkte eines Tiefenbildes von jeweils einem oder mehreren unterschiedlichen Beobachtungspunkten aus. Hierzu werden auch Systeme gezählt, die für (nicht notwendigerweise zusammenhängende) eindimensionale Teilbereiche des Bildes identische Beobachtungspunkte verwenden, zum Beispiel bildspaltenweise. In Abschnitt 2.3.3 werden hierzu Beispiele genannt.

Außerdem wird zwischen *optisch aktiven* und *optisch passiven* Systemkomponenten unterschieden. Die optisch aktiven Komponenten zeichnen sich dadurch aus, daß sie Signale im optischen Frequenzbereich in die beobachtete Szene aussenden. Im Unterschied hierzu bezeichnet der Begriff *geometrisch aktiv* eine Systemkomponente, deren relative Lage zu den übrigen Komponenten vorab durch Kalibrierung bekannt ist und für die Meßpunkterfassung genutzt wird.

2.3.1 Unifokale Systeme

Zu den unifokalen Meßprinzipien zählen die Lichtlaufzeitmessung, die Fokussierungsmessung sowie die Fresnel-Diffraktion.

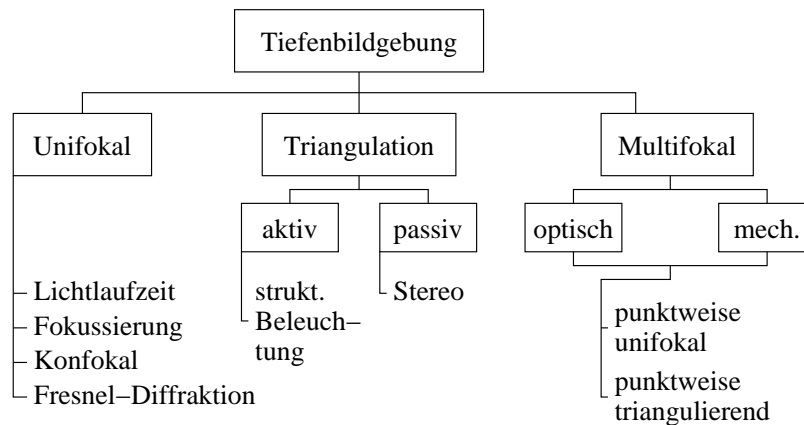


Abbildung 2.1: Klassifikation tiefenbildgebender Verfahren nach Anzahl der beteiligten Beobachtungspunkte.

2.3.1.1 Lichtlaufzeitsysteme

Das Laufzeitprinzip, das auch in zahlreichen anderen Meßverfahren wie beispielsweise der Sonographie, dem Echolot und dem Radar Anwendung findet, basiert auf der einfachen Beziehung

$$c \cdot \delta t = 2r \quad (2.1)$$

mit der Ausbreitungsgeschwindigkeit c des Signals im jeweiligen Medium (Lichtgeschwindigkeit), der gemessenen Zeit δt vom Aussendezeitpunkt bis zum Detektionszeitpunkt nach Reflektion von der Objektoberfläche, und der gemessenen Entfernung r . Entscheidend für die Vermeßbarkeit einer Oberfläche ist die Stärke des reflektierten Signals in Richtung des Detektors, die eine eindeutige Unterscheidung vom Rauschen der Umgebung zulassen muß. Nach Besl [7] verhält sich die Stärke des Signals am Detektor P_R gemäß der Formel

$$P_R = K(t - \delta t) \rho \cos \theta / r^2 \quad (2.2)$$

in der r die Distanz, θ den Winkel zwischen Oberflächennormale und Meßrichtung, ρ einen Koeffizienten der diffusen Oberflächenreflektion, und K eine Funktion bezeichnet, die diverse gerätespezifische Einflußgrößen berücksichtigt. Mit Blick auf die automatisierte Oberflächenvermessung kann also eine schlechtere Datenqualität erwartet werden, wenn die Oberfläche aus einem flacheren Winkel oder aus größerer Entfernung vermessen wird.

Praktisch wird das Signal typischerweise von einem Laser erzeugt, der sequentiell in zwei Richtungen mechanisch abgelenkt wird und so ein Rasterbild erzeugt. Für kurze Distanzen werden neuerdings auch Leuchtdiodenfelder eingesetzt, die das komplette

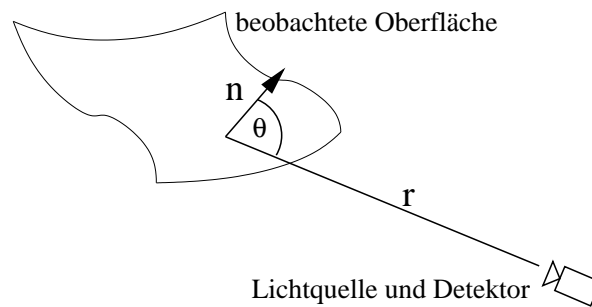


Abbildung 2.2: Konfiguration von Sensor und Oberflächenpunkt bei der Lichtlaufzeitmessung.

Blickfeld parallel beleuchten und so die zeitsynchrone Erfassung eines vollständigen Tiefenbildes ermöglichen [116, 81].

Zur Bestimmung der Lichtlaufzeit werden unterschiedliche Verfahren eingesetzt:

Puls-Detektion: Hier wird ein einfacher Lichtpuls erzeugt und seine Laufzeit gemessen.

Amplituden-Modulation: Durch Änderung der an der Laser-Diode angelegten Spannung wird die Amplitude des erzeugten Lasersignals moduliert. Die Laufzeit kann durch Messung der Phasendifferenz kontinuierlich bestimmt werden, allerdings treten an den Grenzen des Eindeutigkeitsbereichs Phasensprünge auf.

Frequenz-Modulation: Durch Modulation der an der Laser-Diode angelegten Spannung kann das erzeugte Laserlicht in seiner Frequenz (Wellenlänge) moduliert werden und ebenfalls durch Signalvergleich mit dem Referenz-Signal eine Entfernungsbestimmung erreicht werden. Auf die Details des Meßprinzips soll hier nicht näher eingegangen werden.

Die Lichtlaufzeitmessung ist ein optisch aktives Verfahren, daß je nach Klassifizierung des eingesetzten Lasers unter Umständen Maßnahmen zum Augenschutz (Schutzbrille, Sperrung des Meßbereichs für Unbeteiligte) erfordert. Mit kommerziell erhältlichen Systemen lassen sich Meßbereiche von über 150 Metern abdecken.

Zur Erzeugung eines zweidimensionalen Bildrasters muß die nativ punktweise messende Laser-Detektor-Kombination in eine Ablenkeinheit integriert werden. Diese besteht typischerweise aus einem rotierenden Prismenspiegel zur horizontalen Ablenkung, sowie einem weiteren rotierbaren Spiegel zur vertikalen Ablenkung (Abb. 2.3).

Die Geometrie des Bildkoordinatensystems eines derart aufgebauten Sensors entspricht näherungsweise einem Polarkoordinatensystem, so daß sich die gemessene 3D-Koordinate mit der gemessenen Entfernung r und den beiden Ablenkwinkeln ϑ

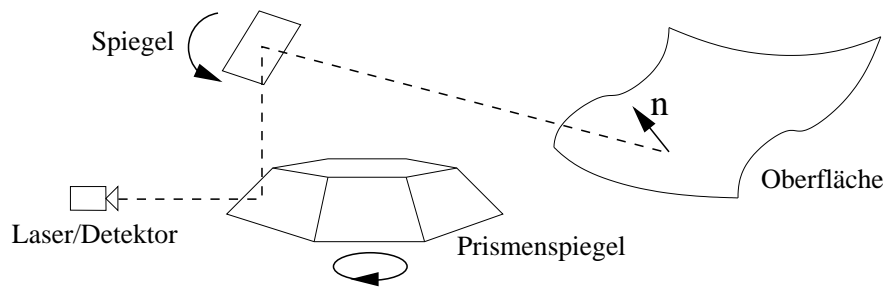


Abbildung 2.3: Erzeugung eines zweidimensionalen Bildes mit punktwiser Abtastung und horizontaler und vertikaler Ablenkung über Spiegel.

und ϕ durch

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = r \begin{pmatrix} \cos \vartheta \cos \varphi \\ \sin \vartheta \\ \cos \vartheta \sin \varphi \end{pmatrix} \quad (2.3)$$

beschreiben läßt. Tatsächlich muß in dieser Vorschrift zusätzlich der Offset zwischen den Ablenkachsen berücksichtigt werden. Außerdem müssen etwaige Ungenauigkeiten des mechanischen Aufbaus modelliert und korrigiert werden. Gerade letzteres ist oftmals nicht systematisch möglich, so daß die Fehlerreduktion in der Praxis durch Tabellen erfolgt, deren Inhalt bei der Kalibrierung bestimmt wird. Abb. 2.4 zeigt eine derartige Korrekturwerttabelle, in der für jeden horizontalen Winkel bildspaltenweise die Abweichung von den Idealkoordinaten (äquidistante Winkelschritte zwischen den Bildspalten) aufgetragen ist.

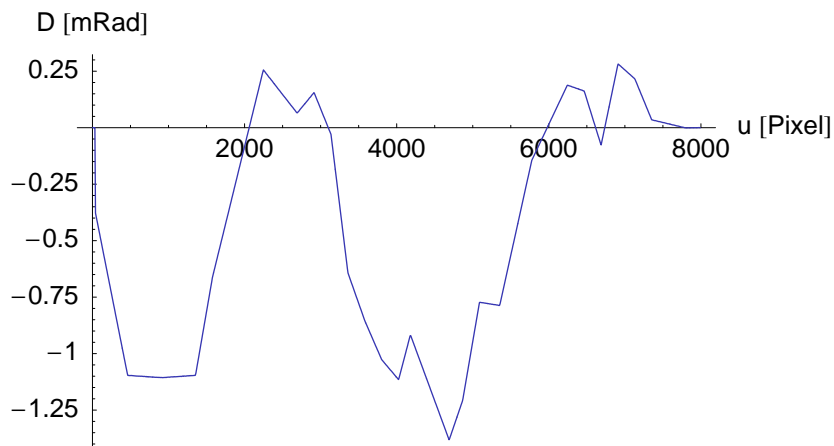


Abbildung 2.4: Einträge der Korrekturwerttabelle der Winkelabweichungen der horizontalen Ablenkeinheit des K2T-Scanners für jede Bildspalte.

2.3.1.2 Fokussierung

Wird ein punktförmiges Merkmal mit einer Kamera durch ein Linsensystem betrachtet, so wird dieses auf der Bildebene nur dann auf einen Punkt abgebildet, wenn das Gauss'sche Gesetz dünner Linsen erfüllt ist (anschaulich in Abbildung 2.5 illustriert). Für einen Punkt in Entfernung z vom Zentrum des Linsensystems mit Brennweite f ist dies der Fall, wenn der Abstand w der Bildebene vom Zentrum des Linsensystems die Bedingung

$$\frac{1}{w} + \frac{1}{z} = \frac{1}{f} \quad (2.4)$$

erfüllt. Punktförmige Merkmale, für die diese Bedingung nicht erfüllt ist, werden aufgrund der Diffraktion an der Blendenöffnung unscharf abgebildet. Aus dem Unschärferadius der Abbildung eines Punktes kann nun seine vorzeichenlose Distanz in z -Richtung zur fokussierten Ebene bestimmt werden, indem zunächst die Intensitätsverteilung des Punktmerkmals im Bild durch eine 2D-Gauss-Verteilung mit dem Durchmesser σ abgeschätzt wird. Der z -Wert ergibt sich bei diesem Ansatz durch

$$z_{\pm} = \frac{wf}{w - f \pm \sigma F} \quad (2.5)$$

wobei $F = f/D$ aus dem Blendenöffnungsdurchmesser D abgeleitet wird (Darstellung nach Besl [7]).

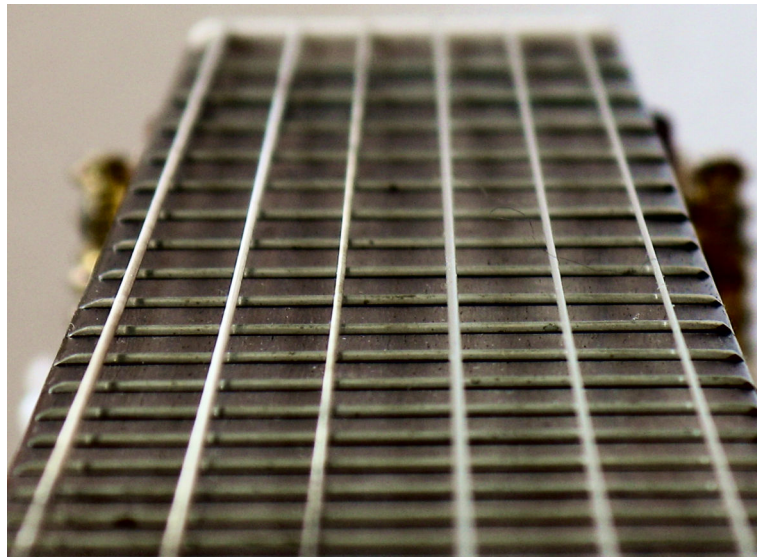


Abbildung 2.5: Die Fokuslage kann im Bild aus dem Unschärferadius bekannter Geometrien oder aus dem lokalen Kontrast bestimmt und zur Entfernungsmessung genutzt werden.

Aus diesen Zusammenhängen kann auf unterschiedliche Weise die Oberflächengeometrie bestimmt werden. Durch Projektion von Punkt- oder Linienmustern kann entweder direkt aus dem Unschärfedurchmesser im Bild bis auf das Vorzeichen der jeweilige Abstand bestimmt werden, oder der Wert w wird mechanisch durch unterschiedliche Fokussierung der Linse variiert, so daß für jeden Punkt im Bild die Distanz durch Minimierung des Unschärfedurchmessers bestimmt werden kann.

Nayar et al. [70] verwenden beispielsweise das Grundprinzip der Fokussierung, um einen in Video-Echtzeit arbeitenden Tiefenbild-Sensor aufzubauen. Hierzu werden im gleichen Strahlengang ein Musterprojektor und zwei Kameras mit unterschiedlicher Fokussierung angeordnet, so daß durch Vergleich der beiden Kamerabilder die Unschärfe der projizierten Muster quantifiziert und daraus die Entfernung abgeleitet werden kann².

Rohály und Hart [91] schlagen einen Ansatz vor, dessen Grundprinzip zwischen Fokussierung und Triangulation angesiedelt ist. Anstelle einer Blende, die durch eine große Blendenöffnung für Verunschärfung sorgt, beschränken sie den Strahlengang durch eine kleine Blendenöffnung, die neben der Hauptachse angeordnet ist und mechanisch um die Hauptachse rotiert werden kann. Mit der Bewegung der Blende wird ein punktförmiges Merkmal auf einen Kreis im Bild abgebildet, dessen Radius von der zu messenden Distanz abhängt. Der Zusammenhang zwischen dem gemessenen Radius d , dem Radius R der rotierten Blendenposition relativ zur Hauptachse, der Brennweite f , der fokussierten Distanz L und dem gesuchten Entfernungswert z ergibt sich durch

$$z = \left(\frac{1}{L} + d \cdot \frac{L - f}{2RfL} \right)^{-1} \quad (2.6)$$

Um den Radius im Bild zu bestimmen, sind wenigstens zwei um 180° versetzte Einzelaufnahmen auszuwerten. Ein ähnlicher Ansatz wird auch von Loranger et al. [60] vorgeschlagen.

2.3.1.3 Konfokale Entfernungsmessung

Im Unterschied zur oben beschriebenen Entfernungsmessung durch Fokussierung wird bei der konfokalen Mikroskopie [68] die Auswertung der Fokussierung nicht durch Bildanalyse, sondern bereits durch den optischen Aufbau erreicht. Hier wird das Licht am Ende des Strahlengangs durch eine sehr kleine Blende geführt, so daß nur Licht aus der Fokusebene den Detektor erreicht. Es handelt sich also um ein optisches Schnittbildverfahren. Durch Variation der Fokussierung oder des z -Abstands kann eine Schnittbildsequenz für unterschiedliche z -Ebenen erzeugt und aus dieser die Oberflächengeometrie rekonstruiert werden.

²Durch den Einsatz telezentrischer Optiken ist das beschriebene Gesamtsystem eigentlich den multifokalen Systemen zuzurechnen.

Der mechanische Scanvorgang kann durch Codierung des gemessenen Lichts vermieden werden. Tiziani und Uhde schlagen dazu das Verfahren der chromatisch-konfokalen Entfernungsmessung [108] vor, bei dem eine wellenlängenabhängige Fokussierung durch eine geeignete Linse erfolgt.

2.3.1.4 Fresnel-Diffraktion

Kohärentes Licht, das durch ein optisches Gitter fällt, kann durch Ausnutzung des Talbot-Effekts (interferometrische Selbstabbildung des optischen Gitters) zur Entfernungsmessung genutzt werden [27]. Wenn die Gitter-Periodizität p wesentlich größer als die Wellenlänge λ des Lichts ist, so treten für Ganzzahlen n in regelmäßigen Abständen $D(n) = 2np^2/\lambda$ scharfe Abbildungen des Gitters auf. Zusätzlich treten in Abständen $D'(n) = (2n + 1)p^2/\lambda$ um 180 Grad phasenverschobene scharfe Abbildungen des Gitters auf. Dazwischen wird das Gitter auf berechenbare Weise unscharf abgebildet, so daß aus dem lokalen Kontrast im Bild innerhalb eines Eindeutigkeitsbereichs die Distanz abgeleitet werden kann. Der typische Aufbau besteht neben dem optischen Gitter aus einer kohärenten Lichtquelle, aus deren Strahlengang mit einem halbdurchlässigen Spiegel das Kamerabild entnommen wird. Nach Besl [7] kann mit diesem Ansatz aufgrund einer Schwierigkeit der exakten Bestimmung des lokalen Kontrasts keine hohe Auflösung erreicht werden.

2.3.2 Triangulations-basierte Systeme

Im Unterschied zu den im vorigen Abschnitt beschriebenen unifokalen Meßverfahren werden bei den triangulations-basierten Verfahren zumindest zwei signifikant auseinanderliegende Beobachtungspunkte benötigt, von denen im Regelfall der eine optisch aktiv, der andere optisch passiv ist (Projektor-Kamera-Kombination). Zu den triangulations-basierten Meßprinzipien zählen im weiteren Sinne auch die Moiré-Interferometrie und die holographische Interferometrie.

2.3.2.1 Meßprinzip

Das Grundprinzip der Entfernungsmessung durch Triangulation ist seit der Antike bekannt. Es existieren nachvollziehbare Beschreibungen, auf welche Weise Thales von Milet im 6. Jahrhundert v. Chr. die Entfernung von Schiffen mit dem Prinzip der Triangulation gemessen hat [49] (vermutlich unter Verwendung tabellierter Kotangens-Werte). Zu diesem Zeitpunkt war die Problemstellung der Kotangens-Bestimmung bereits seit weiteren 1000 Jahren aus der ägyptischen Mathematik überliefert [11].

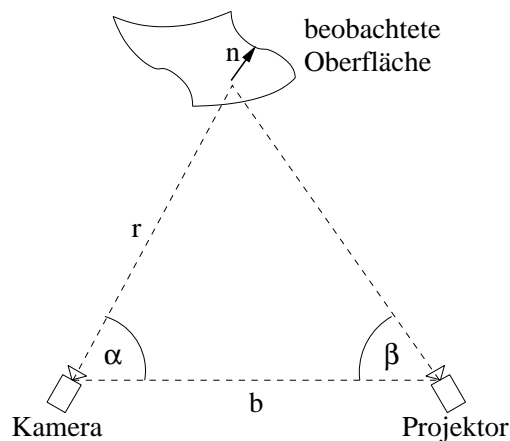


Abbildung 2.6: Konfiguration von Sensor und Oberflächenpunkt bei der Entfernungsmessung durch Triangulation.

Aus der Kenntnis einer Länge und zweier Winkel im Dreieck können die übrigen Größen abgeleitet werden. Während die bekannte Länge, die sogenannte Basislinie (die Verbindungslinie der beiden Beobachtungspunkte, in Abb. 2.6 mit b bezeichnet), konstant ist, ändern sich die Winkel α und β , unter denen ein Punkt jeweils beobachtet wird. Die gesuchte Entfernung ergibt sich durch

$$r = b \cdot \frac{\sin \beta}{\sin \vartheta} \quad (2.7)$$

mit dem Triangulationswinkel $\vartheta := \pi - \alpha - \beta$. Zusätzlich ist für die Bestimmung der gemessenen 3D-Position noch ein weiterer Winkel relevant, der die Drehung des von Meßpunkt und Beobachtungspunkten aufgespannten Dreiecks um die Basislinie beschreibt.

Im Kontext kamerabasierter Sensoren wird die zugrundeliegende Geometrie der zwei Blickpunkte durch das Epipolarbündel beschrieben, das aus der Menge aller Ebenen besteht, die die Basislinie enthalten, den sogenannten Epipolar-Ebenen. Abbildung 2.7 zeigt die entsprechende Konfiguration der zwei Beobachtungspunkte c und c' , vor denen sich jeweils die Bildebenen der Kamera befinden; die Epipolar-Ebene ist grau dargestellt. Der Schnittpunkt der jeweiligen Bildebene mit der Basislinie wird als Epipol bezeichnet (im Bild mit e beziehungsweise e' bezeichnet), die Schnittkurven von Bildebene und Epipolar-Ebenen heißen Epipolar-Kurven (wegen der nicht-linearen perspektivischen Abbildung realer Linsensysteme aufgrund der Linsenverzerrung existieren die sogenannten Epipolar-Geraden im Allgemeinen nur in nachträglich entzerrten Bildern).

Offensichtlich wird die zu einem beobachteten Punkt p gehörige Epipolar-Ebene bereits durch die Lage dieses Punktes im Bild der ersten Kamera (in c) bestimmt, sofern

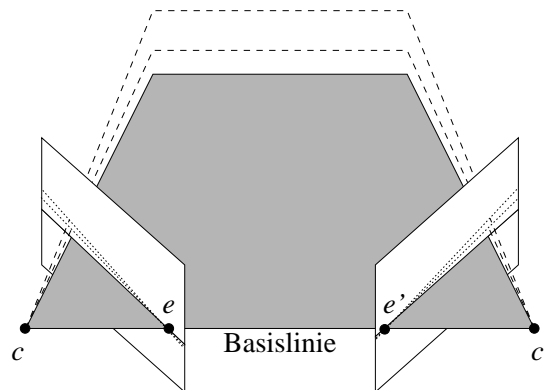


Abbildung 2.7: *Epipolar-Geometrie zweier Blickpunkte c und c' mit Epipolen e, e' in den Bildebenen sowie dem Epipolarbündel (grau bzw gestrichelt, Darstellung in Anlehnung an [43]).*

der beobachtete Punkt nicht auf den Epipol abgebildet wird (in diesem Fall würde die Triangulation auf einem degenerierten Dreieck basieren und ist daher nicht möglich). Der zweite Beobachtungspunkt muß nun lediglich die eindimensionale Information liefern, unter welchem Winkel der Punkt p von c' aus relativ zur Basislinie beobachtet wird. Im Bild der Kamera in c' sind die möglichen Lagen auf die Epipolar-Kurve beschränkt.



Abbildung 2.8: *Szene unter strukturierter Beleuchtung durch einen Streifenlichtprojektor.*

Um auf der Epipolar-Geometrie ein flächenhaftes Meßverfahren aufzubauen, müssen daher zwei Teilprobleme gelöst werden. Zunächst ist die Kamerageometrie durch

Kalibrierung zu bestimmen, so daß jedem Punkt in der Bildebene ein Sichtstrahl zugeordnet werden kann. Die Kamerageometrie besteht aus der extrinsischen Geometrie, die die relative Lage der Kameras zueinander beschreibt, sowie der intrinsischen Geometrie, die die Abbildung eines Punktes durch das Linsensystem auf das in der Bildebene liegende Bildraster beschreibt. Optisch aktive Beobachtungspunkte (Projektoren) unterscheiden sich in den Anforderungen zunächst nicht von optisch passiven Beobachtungspunkten (Kameras), da auch für sie zu jedem Punkt im projizierten Muster der Sichtstrahl bekannt sein muß, um eine Triangulation durchführen zu können.

Aufgrund der Epipolar-Geometrie genügt es allerdings für den zweiten Beobachtungspunkt c' , die Kamera- beziehungsweise Projektor-Geometrie nur insoweit zu bestimmen, daß nicht der Sichtstrahl, sondern lediglich eine den Sichtstrahl enthaltende Sichtfläche zu jedem Bildpunkt bekannt ist, welche näherungsweise senkrecht zur Epipolar-Ebene angeordnet ist (zum Beispiel Linien-, Kurven- oder Streifenmusterprojektion). Die Kenntnis der Lage des Sichtstrahls innerhalb der Sichtfläche wird für c' zur Triangulation nicht benötigt.

Das zweite der Teilprobleme ist die Bestimmung der Korrespondenzpunkte in der Bildebene der Kamera am zweiten Beobachtungspunkt c' . Die in der Literatur beschriebenen Verfahren zur passiven Rekonstruktion der Oberflächengeometrie aus Stereobildpaaren [47, 93, 82, 59] setzen eine signifikante Texturierung der Oberfläche voraus und sind daher für einen großen Teil möglicher Rekonstruktionsobjekte ungeeignet. Daher werden die Ebenen des Epipolar-Bündels im Regelfall durch Einsatz von Musterprojektion strukturiert (Abbildung 2.8), so daß von c aus beobachteten Punkten auf der Objektoberfläche eindeutig ein korrespondierender Punkt auf der jeweiligen Epipolar-Kurve im Bild von c' zugeordnet werden kann.

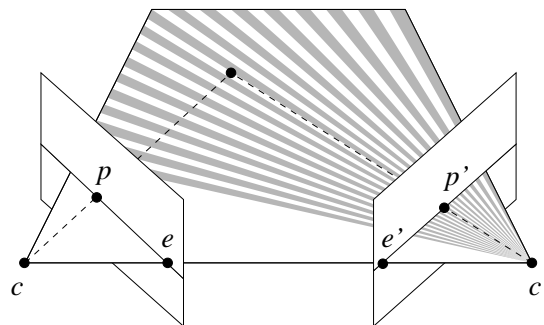


Abbildung 2.9: Aktive Musterprojektion, bei der die Geometrie des Projektors bekannt ist und in die Triangulation einbezogen wird.

Dies kann entweder, wie bereits oben angedeutet, durch ein kalibriertes Projektionsgerät am zweiten Beobachtungspunkt realisiert werden. Die Winkelinformation zu c' wird in diesem Fall auf der Objektoberfläche dargestellt und kann von einer Kamera

in c direkt jedem Bildpunkt zugeordnet werden (Abb. 2.9). Alternativ kann ein geometrisch passiver Projektor eingesetzt werden, der über seine strukturierende Wirkung hinaus keinen Beitrag zur Triangulation leistet und der daher nicht kalibriert werden muß. Die Triangulation wird in diesem Fall über zwei kalibrierte Kameras in c und c' durchgeführt (Abb. 2.10).

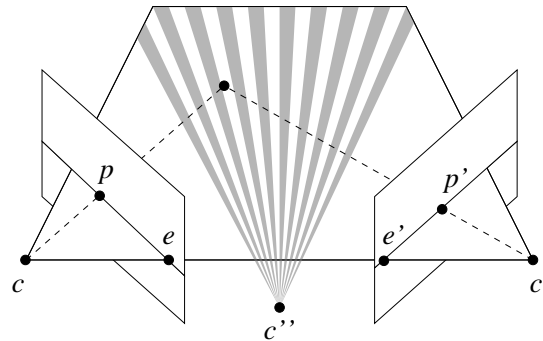


Abbildung 2.10: Passive Musterprojektion, bei der die Triangulierung über zwei Kameras ohne Einbeziehung der Projektorgeometrie erfolgt.

Für die aktive Musterprojektion spricht vor allem der erheblich geringere Rechenaufwand, der darin begründet ist, daß die Korrespondenzpunkte nicht gesucht werden müssen. Der Vorteil der passiven Musterprojektion liegt einerseits im höheren Genauigkeitspotential der Kalibrierung zum zweiten Beobachtungspunkt c' , da hier das Kamerabild direkt ausgewertet werden kann. Im Vergleich dazu müssen zur Projektor-Kalibrierung die Aufnahmen einer mit einem gewissen Restfehler kalibrierten Kamera herangezogen werden, so daß eine Fehlerakkumulation unterstellt werden darf. Zum anderen stellt die passive Musterprojektion nur geringe Anforderungen an die Projektorqualität, so daß kostengünstige Geräte eingesetzt werden können (zum Beispiel Video-Projektoren).

2.3.2.2 Mustercodierung

Ziel der Mustercodierung ist im Falle der passiven Musterprojektion, eine eindeutige Zuordnung korrespondierender Bildpunkte zu ermöglichen. Im Falle der aktiven Musterprojektion ist das Ziel die Rekonstruktion des Projektionswinkels zu jedem Punkt des Musters. Einfache Ansätze verwenden beispielsweise eine einzelne, mit einer Zylinderoptik erzeugte Laserlinie, die durch das Meßfeld geführt wird und für die im Falle der aktiven Musterprojektion der jeweilige Winkel bekannt ist [90, als Beispiel]. Der gravierende Nachteil dieses Verfahrens ist die Notwendigkeit der bildspalten-sequentiellen Abtastung: für jeden Winkel, der mit der Laserlinie unterschieden werden soll, wird eine eigene Kameraaufnahme benötigt. Durch die Verwendung von kohärentem

Laserlicht im Zusammenspiel mit schmalbandigen Detektorfiltern kann jedoch eine hohe Robustheit der Musterdetektion erreicht werden, die im industriellen Kontext oftmals entscheidend ist.

Optimal in Bezug auf die Akquisitionszeit ist dagegen die passive Projektion eines statischen Zufallsmusters, das auch in uniform texturierten Oberflächenbereichen eine robuste Korrespondenzetablierung mit nur einer Stereoaufnahme erlaubt. Nach diesem von Schewe und Förstner [94] vorgeschlagenen Verfahren erfolgt die Bestimmung der Korrespondenzpunkte durch Kleinste-Quadrate-Anpassung zweier 8×8 -Fenster im Stereobildpaar. Das Verfahren ist dadurch sehr rechenintensiv.

In der vorliegenden Arbeit wurden Verfahren der Musterprojektion eingesetzt, die die Punkte innerhalb des Projektorbildes durch Kombination mehrerer Einzelmuster unterscheidbar machen, so daß ein streng monoton mit dem horizontalen Bildwinkel steigender Zahlenwert codiert wird. Hierzu werden Phasenschiebeverfahren eingesetzt, die Bilder eines näherungsweise sinusförmig modulierten Streifenmusters in N unterschiedlichen Phasenlagen kombinieren. Mit der Phasenverschiebung $\Phi_n = (n - 1)(2\pi/N)$ kann die beobachtete Intensität zu einem Punkt (x, y) durch

$$I_n(x, y) = a(x, y) + b(x, y) \cdot \cos(\Phi_{obj}(x, y) + \Phi_n) \quad (2.8)$$

dargestellt werden, wobei $a(x, y)$ die Umgebungsbeleuchtung, $b(x, y)$ den Kontrast des projizierten Musters und $\Phi_{obj}(x, y)$ den zu rekonstruierenden Phasenwert bezeichnet. Da die Beziehung in Gleichung 2.8 drei Unbekannte, nämlich neben dem gesuchten Phasenwert $\Phi_{obj}(x, y)$ die Werte $a(x, y)$ und $b(x, y)$ enthält, müssen mindestens drei dieser Gleichungen für unterschiedliche n aufgestellt werden. Aus diesen kann der gesuchte Phasenwert durch

$$\tan \Phi_{obj}(x, y) = \frac{\sum_{n=1}^N I_n(x, y) \sin(2\pi n/N)}{\sum_{n=1}^N I_n(x, y) \cos(2\pi n/N)} \quad (2.9)$$

berechnet werden. Für $N = 4$ mit einem Phasen-Versatz von $\frac{\pi}{2}$ ergibt sich beispielsweise die einfache Formel

$$\tan \Phi_{obj}(x, y) = \frac{I_1(x, y) - I_3(x, y)}{I_4(x, y) - I_2(x, y)} \quad (2.10)$$

Durch die Periodizität des projizierten Sinusmusters kann auf diese Weise allerdings nur der Abschnitt innerhalb einer Sinuswelle, nicht aber die Wellenzahl codiert werden, so daß auch die absolute horizontale Lage innerhalb des Projektorbildes nicht bestimmt werden kann. Zur Codierung der Wellenzahl wird meist ein binärer Gray-Code eingesetzt [92] (in Abbildung 2.11 dargestellt), alternativ kommen auch Mehrwellenlängenphasenschiebeverfahren [84, 85, 38] in Frage.

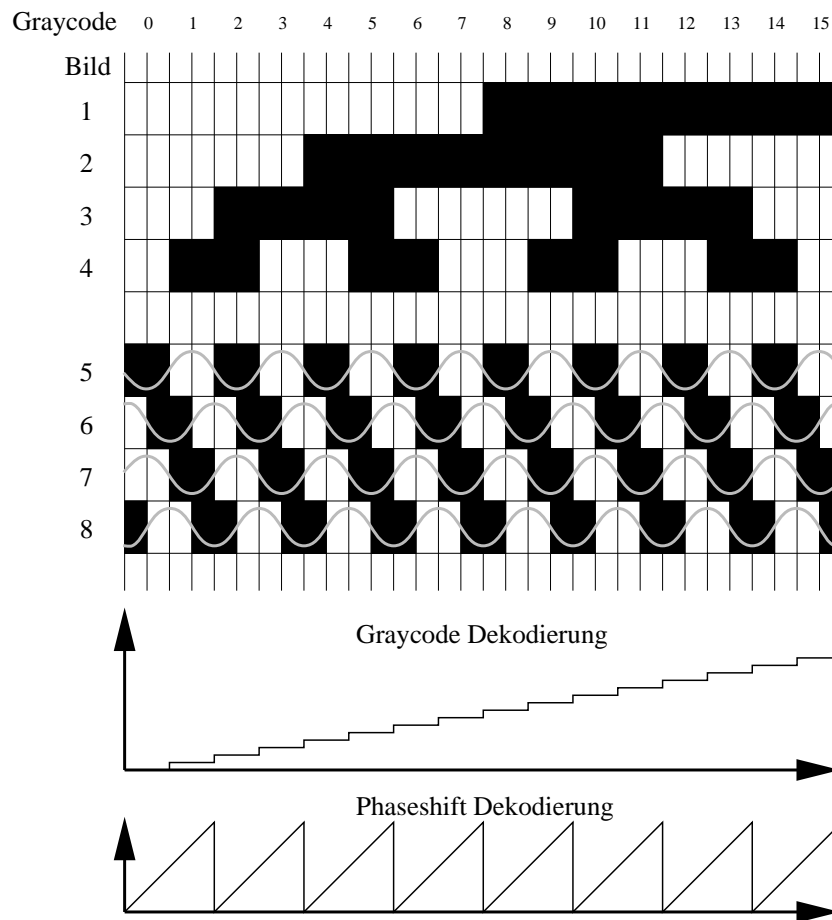


Abbildung 2.11: Kombination von Gray-Code und Phaseschiebeverfahren zur Bestimmung von Phase und Wellenzahl.

2.3.2.3 Kameramodell und Kalibrierung

Die Geometrie handelsüblicher Flächenkameras und ihrer Linsensysteme wird im Rahmen der vorliegenden Arbeit durch eine Lochkamera mit punktförmiger Blendenöffnung modelliert; zusätzlich wird die Verzeichnung des Linsensystems modelliert. Mathematisch läßt sich damit die Abbildung eines in Weltkoordinaten vorliegenden Punktes p_w auf das Bildraster in den folgenden Schritten beschreiben. Mit den Parametern der extrinsischen Kamerageometrie wird der Punkt zunächst durch Rotation R und Translation t in das Koordinatensystem der Kamera transformiert:

$$P_c = R \cdot p_w + t \quad (2.11)$$

Per Konvention koinzidiert die positive z -Achse des Kamera-Koordinatensystems mit der optischen Achse der Kamera, während der Ursprung des Kamera-Koordinatensystems

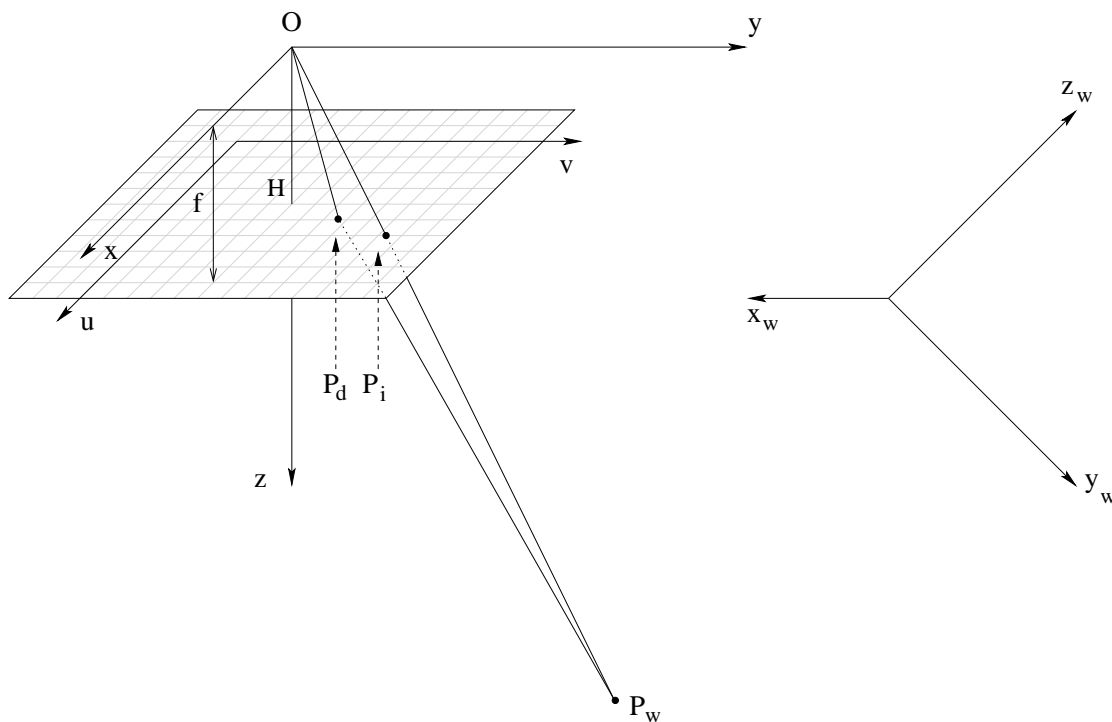


Abbildung 2.12: Kamerageometrie mit Weltkoordinatenachsen (x_w, y_w, z_w) , Kamerakoordinatenachsen (x, y, z) und Bildkoordinatenachsen (u, v) . Aufgrund der Linsenverzeichnung ergibt sich ein zu korrigierender Versatz zwischen den idealen Rasterkoordinaten P_i und den tatsächlichen Rasterkoordinaten P_d des projizierten Punktes P_w .

stems im optischen Zentrum liegt, wie in Abbildung 2.12 dargestellt. Mit den Parametern der intrinsischen Kamerageometrie wird der nun in Kamerakoordinaten vorliegende Punkt $p_c = (x, y, z)^T$ auf das Bildraster durch

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} \frac{\alpha x}{z} + u_0 - \Delta_u \\ s \frac{\alpha y}{z} + v_0 - \Delta_v \\ g(z) \end{pmatrix} \quad (2.12)$$

abgebildet. Dabei ist (u, v) die Position im Bildraster, und $(u_0, v_0) = H$ der Bildhauptpunkt, an dem die optische Achse die Bildebene schneidet. Der Abstand d der Bildebene vom optischen Zentrum, der bei Fokussierung auf Unendlich der Brennweite f entspricht, geht zusammen mit der Bildrasterskalierung durch den Wert α in die Rechnung ein, so daß die tatsächliche Chipgröße und die Brennweite rechnerisch nicht getrennt werden müssen, sondern nur das Pixelseitenverhältnis s berücksichtigt wird. Durch Δ_u, Δ_v werden Abbildungsfehler berücksichtigt, die durch radial-symmetrische Verzeichnung und durch radial-asymmetrische und tangentielle Verzeichnung zustande-

kommen. Vollständigkeitshalber und um die Definition der Umkehrfunktion zu ermöglichen, wird auch der an der Rasterposition (u, v) vorliegende z -Wert einem im Bildraaster zu repräsentierenden w -Wert zugeordnet. Dies geschieht durch eine beliebige stetige und bijektive Funktion $g : \mathbb{R} \rightarrow \mathbb{R}$, die typischerweise den Erfordernissen der Zahlenrepräsentation Rechnung trägt (zum Beispiel 16bit-Ganzzahl-Repräsentation).

Die Kalibrierung, also die Bestimmung der Parameter, eines derartigen Systems kann im Prinzip als Stand der Technik betrachtet werden [13, 109, 103, 95]. Obwohl aber im Manual of Photogrammetry [101, zitiert nach [102]] vor der starken Korrelation der einzelnen Modellparameter als einem der zentralen Probleme der Kamerakalibrierung gewarnt wird, sind einige veröffentlichte Verfahren diesbezüglich dennoch von Heuristiken (zum Beispiel [113]) geprägt, sofern die Problematik überhaupt thematisiert wird. Daher werden im Folgenden eigene Arbeiten zur systematischen Dekorrelation der Kameraparameter beschrieben, welche die numerische Optimierung entscheidend vereinfachen. Zur Gewinnung der Näherungslösung wird dagegen auf die obengenannte Literatur verwiesen.

Die in den Gleichungen 2.11 und 2.12 beschriebene Abbildung weist Korrelationen zwischen den Translationsparametern t_x, t_y mit der Position des Bildhauptpunktes u_0, v_0 sowie zwischen dem Translationsparameter t_z und der Kamerakonstanten α auf, weiterhin bestehen Zusammenhänge zwischen diesen Parametern und den Koeffizienten der radialen und tangentialen Linsenverzeichnung (in Δ_u, Δ_v zusammengefaßt). Änderungen in einem der intrinsischen Parameter müssen daher im jeweils korrelierenden extrinsischen Parameter kompensiert werden, um den beabsichtigten Effekt der intrinsischen Parameteränderung zu isolieren.

Eine Änderung des Bildhauptpunktes bewirkt beispielsweise eine Translation aller abgebildeten Punkte in der Bildebene. Beabsichtigt ist aber eine Änderung des Zentrums der perspektivischen Abbildung, also der relativen Translationen der abgebildeten Punkte in Abhängigkeit von ihrer Distanz zur optischen Achse. Diese Diskrepanz ist für die numerische Optimierung ungünstig und erfordert den Einsatz rechenaufwändiger Verfahren. Im Folgenden wird daher gezeigt, wie die Parameter rechnerisch dekoriert werden können beziehungsweise Korrelationen durch geeignete Änderungen des Kameramodells vermieden werden können.

Kamerakonstante Anstelle der klassischen Formulierung der perspektivischen Projektion, bei der das optische Zentrum im Koordinatenursprung liegt, und die Bildebene durch $z = d$ definiert wird (Abbildung 2.13 links), ist eine alternative Formulierung [31] numerisch günstiger, bei der die Bildebene durch $z = 0$ definiert ist, und das optische Zentrum entsprechend verschoben an der Stelle $z = -d$ auf der z -Achse liegt (Abbildung 2.13 rechts).

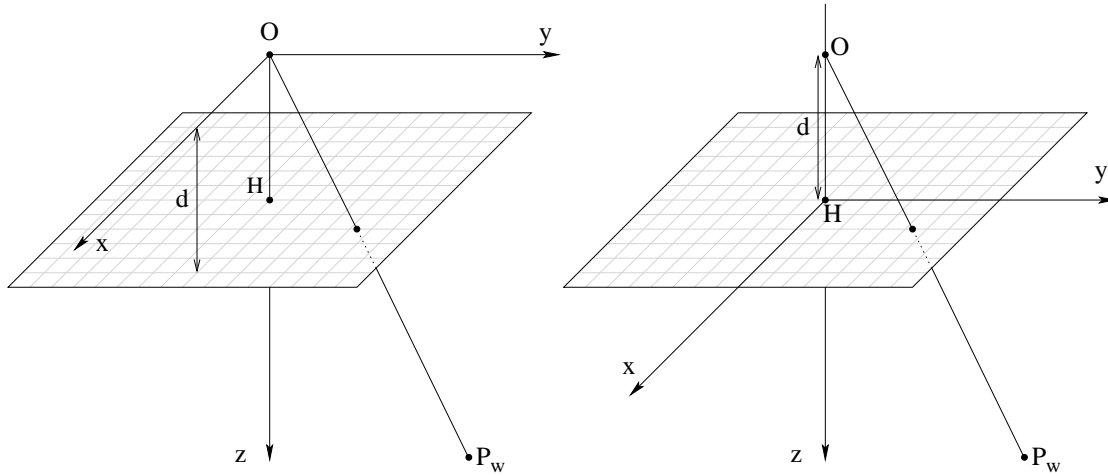


Abbildung 2.13: Klassische (links) und alternative (rechts) Anordnung des Koordinatensystems relativ zu optischem Zentrum und Bildebene.

Rechnerisch ändert sich dadurch die Formulierung der perspektivischen Projektion auf die Bildebene gegenüber Gleichung 2.12 in

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} \frac{x}{z \cdot \alpha^{-1} + 1} + u_0 - \Delta_u \\ s \frac{y}{z \cdot \alpha^{-1} + 1} + v_0 - \Delta_v \\ g(z) \end{pmatrix} \quad (2.13)$$

Im Falle einer nahezu telezentrischen Optik mit $\alpha \rightarrow \infty$ können numerische Probleme so auf elegante Weise vermieden werden ($z \cdot \alpha^{-1} \rightarrow 0$). Die Optimierung von α^{-1} bietet außerdem den Vorteil einer gleichmäßigeren Verteilung der Änderungen im Bildraum in Abhängigkeit von der jeweiligen Parameteränderung im Wertebereich.

Zusätzlich ist es sinnvoll, Änderungen der Kamerakonstante durch eine entsprechende Anpassung der Translation in Richtung der optischen Achse zu kompensieren, so daß sich die Lage der abgebildeten Objektpunkte für die mittlere Objekt-Kamera-Distanz m_z nicht ändert. Jedem $\delta_{\alpha^{-1}}$ muß dabei ein entsprechendes δ_z mit

$$\delta_z = -\frac{\delta_{\alpha^{-1}} \cdot m_z}{\alpha^{-1} + \delta_{\alpha^{-1}}} \quad (2.14)$$

zugeordnet werden, welches bei der Berechnung partieller Ableitungen und der Umsetzung von Korrekturwerten im Rahmen der numerischen Optimierung zur Anwendung kommt.

Bildhauptpunkt Die Parameter der Lage des Bildhauptpunktes $H = (u_0, v_0)$ korrelieren numerisch mit der lateralen Translation parallel zur Bildebene (Parameter t_x und

t_y). Analog zu der in Gleichung 2.14 beschriebenen Vorgehensweise kann auch hier eine Dekorrelation erreicht werden, indem zu jeder Änderung δ_{u_0} beziehungsweise δ_{v_0} eine entsprechende Kompensation in der äußeren Orientierung t_x beziehungsweise t_y vorgenommen wird, so daß sich wiederum die Lage der abgebildeten Objektpunkte für die mittlere Objekt-Kamera-Distanz m_z nicht ändert:

$$\begin{pmatrix} \delta_{t_x} \\ \delta_{t_y} \end{pmatrix} = \begin{pmatrix} -\delta_{u_0}(1 + m_z \cdot \alpha^{-1}) \\ -\frac{1}{s}\delta_{v_0}(1 + m_z \cdot \alpha^{-1}) \end{pmatrix} \quad (2.15)$$

Auch diese Kompensationsterme kommen bei der Berechnung partieller Ableitungen und der Umsetzung von Korrekturwerten im Rahmen der numerischen Optimierung zur Anwendung.

Radial-symmetrische Verzeichnung Die radial-symmetrische Verzeichnung ist eine Variation des Abbildungsmaßstabs in Abhängigkeit vom Bildradius und hängt daher unmittelbar mit der Kamerakonstanten zusammen. Ursächlich hängt die radial-symmetrische Verzeichnung von rotationssymmetrischen Fertigungsfehlern der Linse, aber auch von Fokussierung und Blendenöffnung ab. Die auftretenden Fehler können für handelsübliche Objektive in den Bildecken leicht 1% des Bildradius überschreiten [62] und müssen daher zwingend korrigiert werden.

In der klassischen Formel nach Brown [13] wird die Verzeichnung durch eine Taylor-Reihe mit

$$\begin{pmatrix} \Delta u_{rad} \\ \Delta v_{rad} \end{pmatrix} = \begin{pmatrix} \bar{u}(\kappa_1 r^2 + \kappa_2 r^4 + \kappa_3 r^6 + \dots) \\ \bar{v}(\kappa_1 r^2 + \kappa_2 r^4 + \kappa_3 r^6 + \dots) \end{pmatrix} \quad (2.16)$$

korrigiert, wobei $\bar{u} = u - u_0$, $\bar{v} = v - v_0$ und $r = \sqrt{\bar{u}^2 + \bar{v}^2}$.

Fryer [32, zitiert nach [62]] schlägt zur Beschreibung dieser Verzeichnung einen Polynom-Ansatz vor, bei dem die Verzeichnungskorrektur an einem frei wählbaren Radius r_0 Null ist. r_0 wird dabei so festgelegt, daß die minimal und maximal auftretenden Verzeichnungswerte gleich groß sind (*balanced radial distortion*). Durch die entstehende Symmetrie wird die Korrelation mit der Kamerakonstanten minimiert, als Verzeichnungskorrektur ergibt sich

$$\begin{pmatrix} \Delta u_{rad} \\ \Delta v_{rad} \end{pmatrix} = \begin{pmatrix} \bar{u}(\kappa_1(r^2 - r_0^2) + \kappa_2(r^4 - r_0^4) + \kappa_3(r^6 - r_0^6) + \dots) \\ \bar{v}(\kappa_1(r^2 - r_0^2) + \kappa_2(r^4 - r_0^4) + \kappa_3(r^6 - r_0^6) + \dots) \end{pmatrix} \quad (2.17)$$

In der vorliegenden Arbeit wurde r_0 vereinfacht auf $2/3$ des maximal auftretenden Bildradius festgelegt, da auf diese Weise der Optimalwert für typische Objektive bereits gut angenähert wird. Abbildung 2.14 und 2.15 zeigt die Verteilung und Richtung der Verzeichnungskorrektur im Bild sowie die absoluten Verzeichnungswerte in Pixeln für ein typisches Objektiv.

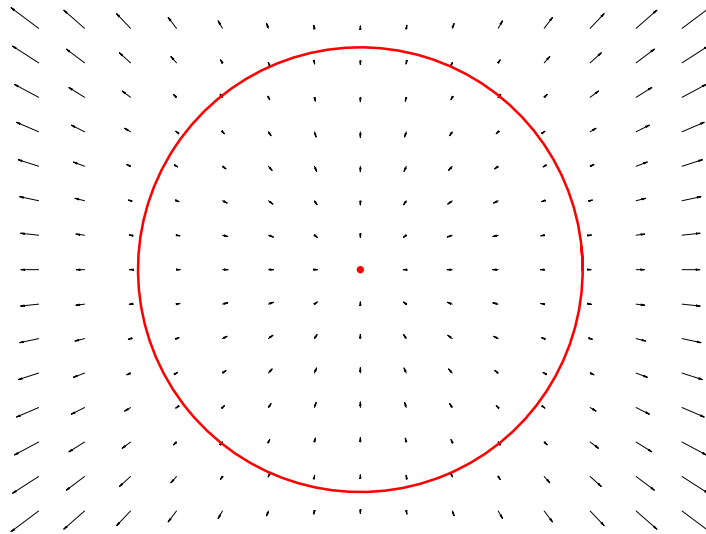


Abbildung 2.14: Korrekturvektoren im Bild für die balancierte radial-symmetrische Linsenverzeichnung.

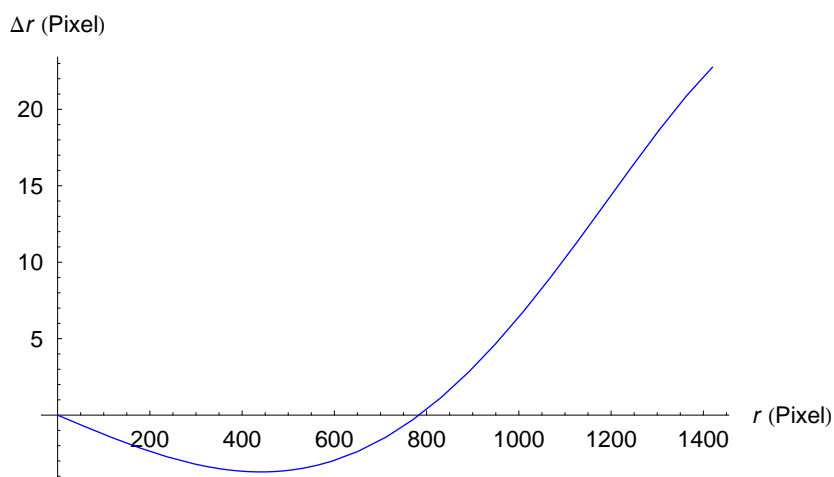


Abbildung 2.15: Korrekturbeträge in Abhängigkeit vom Bildradius für die balancierte radial-symmetrische Linsenverzeichnung.

Radial-asymmetrische und tangentiale Verzeichnung Die radial-asymmetrische und tangentiale Verzeichnung entsteht in erster Linie durch die Dezentrierung der Linsen im Strahlengang des Objektivs. In Abhängigkeit von der Qualität der Objektiv-Mechanik kann sie sich beispielsweise bei der Fokussierung ändern. In der Standardformel nach Conrady [19] wird diese Verzeichnung durch die beiden Parameter b_1, b_2

mit

$$\begin{pmatrix} \Delta u_{tan} \\ \Delta v_{tan} \end{pmatrix} = \begin{pmatrix} b_1(r^2 + 2u^2) + 2b_2uv \\ b_2(r^2 + 2v^2) + 2b_1uv \end{pmatrix} \quad (2.18)$$

modelliert. Wie in Abbildung 2.16 zu sehen, ist die Verzeichnungskorrektur damit in

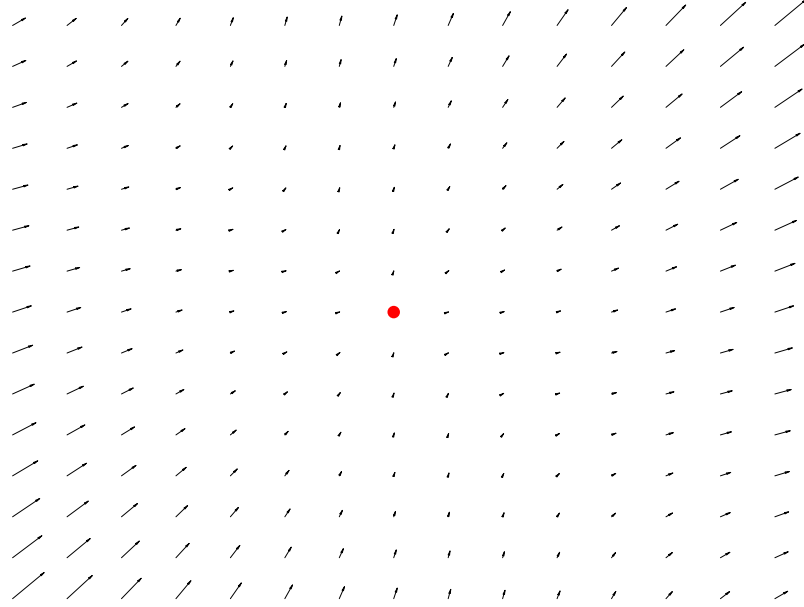


Abbildung 2.16: *Korrekturvektoren der radial-asyymetrischen und tangentialen Verzeichnung im Bild.*

der Bildmitte Null, im Mittel bewirkt die Formel allerdings eine Translation des Bildes, so daß eine Parameterkorrelation mit dem Bildhauptpunkt sowie mit der extrinsischen Translation t_x, t_y zu erwarten ist. Analog zur Einführung eines linearen Anteils bei der Korrektur der radial-symmetrischen Verzeichnung nach Fryer sollen daher die Korrekturvektoren so definiert werden, daß die mittlere Korrektur über der Bildfläche Null ist.

Die mittlere Korrektur ergibt sich durch Integration über die Bildfläche, die hier durch die Parameter u_s, v_s definiert ist, durch

$$\iint_{-u_s, -v_s}^{u_s, v_s} \begin{pmatrix} b_1(3u^2 + v^2) + 2b_2uv \\ b_2(u^2 + 3v^2) + 2b_1uv \end{pmatrix} du dv = \frac{1}{3} \begin{pmatrix} b_1(3u_s^2 + v_s^2) \\ b_2(u_s^2 + 3v_s^2) \end{pmatrix} \quad (2.19)$$

Anstelle der in Gleichung 2.18 beschriebenen Korrektur kann damit eine translationsneutrale Formulierung durch

$$\begin{pmatrix} \Delta u_{tan} \\ \Delta v_{tan} \end{pmatrix} = \begin{pmatrix} b_1(3u^2 + v^2) + 2b_2uv \\ b_2(u^2 + 3v^2) + 2b_1uv \end{pmatrix} - \frac{1}{3} \begin{pmatrix} b_1(3u_s^2 + v_s^2) \\ b_2(u_s^2 + 3v_s^2) \end{pmatrix} \quad (2.20)$$

erreicht werden. In der Praxis wird dabei die Bildgröße u_s, v_s wiederum auf den tatsächlich genutzten Bildbereich beschränkt. Im Ergebnis gibt es für $b_1, b_2 \neq 0$ vier Punkte im Bild, für die die Verzeichnungskorrektur Null ist; diese sind mit Verteilung und Richtung der Korrekturen beispielhaft in Abbildung 2.17 dargestellt.

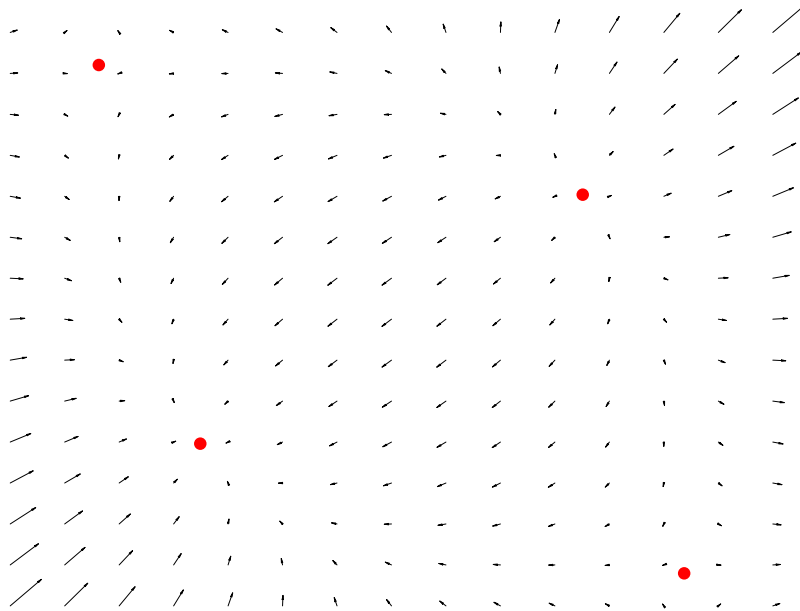


Abbildung 2.17: Korrekturvektoren der radial-asyymmetrischen und tangentialen Verzeichnung nach Subtraktion der mittleren Translation.

2.3.2.4 Fehler der Entfernungsmessung

Die Genauigkeit triangulations-basierter Systeme hängt nicht allein von den Eigenschaften der eingesetzten Geräte, sondern ganz wesentlich auch vom geometrischen Aufbau ab. Traditionell wird davon ausgegangen, daß vor allem die Größe des Triangulationswinkels die erreichbare Meßgenauigkeit bestimmt. Während der laterale Fehler Δ_{lat} allein durch den Winkelfehler ε der Kamera mit $\Delta_{lat} = r \cdot \tan \varepsilon$ abgeschätzt werden kann, ist der Zusammenhang für den Fehler Δ_r der Entfernungsmessung nach Gleichung 2.7 durch den Einfluß mehrerer Fehlerquellen komplizierter. Im Folgenden wird die Berechnung des Meßfehlers beschrieben, und es werden einige für die vorliegende Arbeit relevante Beispielrechnungen durchgeführt.

Zur Analyse des Meßfehlers wird die Entfernungsberechnung als Funktion der beiden Winkel α und β , unter denen der beobachtete Punkt von den beiden Beobachtungspunkten eingesehen wird, sowie der Länge der Basislinie b zwischen den beiden

Beobachtungspunkten durch

$$r(\alpha, \beta, b) := b \cdot \frac{\sin \beta}{\sin(\pi - \alpha - \beta)} \quad (2.21)$$

beschrieben. Da r nach den Variablen α , β und b stetige partielle Ableitungen besitzt, kann die Fehlerschranke nach Bronstein [12] mit Hilfe des totalen Differentials von r näherungsweise durch

$$\Delta r \approx \Delta \alpha \cdot |r_\alpha(\alpha, \beta, b)| + \Delta \beta \cdot |r_\beta(\alpha, \beta, b)| + \Delta b \cdot |r_b(\alpha, \beta, b)| \quad (2.22)$$

angegeben werden. In der Praxis sind die in die Berechnung einfließenden Fehlerquellen $\Delta \alpha$, $\Delta \beta$ und Δb zunächst durch die Qualität der Kalibrierung bestimmt; im laufenden Betrieb werden die Werte zusätzlich durch mechanische Änderungen des Aufbaus wie beispielsweise die temperaturabhängige Ausdehnung des Trägermaterials bestimmt. In der folgenden Beispielrechnung wird von einem symmetrischen Aufbau des Meßsystems und einer Messung in der Mitte des Meßvolumens ausgegangen, so daß $\alpha = \beta = \frac{1}{2}(\pi - \vartheta)$ mit dem Triangulationswinkel ϑ angenommen werden kann. Für einen gegebenen Meßabstand r hängt damit auch die Länge der Basislinie vom Triangulationswinkel über $b = 2r \sin \frac{\vartheta}{2}$ ab. Die drei Fehlersummanden aus Gleichung 2.22 können daher für $\vartheta < \frac{\pi}{2}$ durch

$$\Delta r \approx \Delta \alpha \cdot \frac{r}{\tan \vartheta} + \Delta \beta \cdot \frac{r}{\sin \vartheta} + \Delta b \cdot \frac{\cos \frac{\vartheta}{2}}{\sin \vartheta} \quad (2.23)$$

zur Abschätzung des Entfernungsmeßfehlers zusammengefaßt werden. Zwei einfache Rechenbeispiele sollen dies im Folgenden illustrieren; dabei wird als realistischer Aufbau ein Meßabstand von 1000 mm und ein Stereo-Kamerasystem mit $1/2''$ -CCD-Elementen mit einer Auflösung von 1024×768 Pixeln angenommen, das mit Objektiven der Brennweite $f = 16$ mm ausgerüstet ist (der horizontale Öffnungswinkel von 22.6° ergibt in 1000 mm Abstand eine Sichtfeldbreite von 391 mm).

Zur Abschätzung des Winkelfehlers $\Delta \alpha$ und $\Delta \beta$ wird als Fehlerquelle der Rückprojektionsfehler der Kalibrierung und die Detektionsgenauigkeit des korrespondierenden Punktes auf der Oberfläche (mit strukturierender Musterprojektion) angenommen, die etwa 0.05 Pixel beträgt. Mit der CCD-Größe ($1/2''$ entspricht 6.4×4.8 mm²) ergibt sich daraus über Gleichung 2.12 in der Bildmitte unter Vernachlässigung des Verzeichnungsfehlers ein Winkelfehler von 19.53 μ Rad. Abbildung 2.18 zeigt Δr für diese Konstellation in Abhängigkeit vom Triangulationswinkel ϑ und unter der Annahme $\Delta b = 0$.

Eine weitere, im Kontext der vorliegenden Arbeit wichtige Fehlerquelle ist die elastische Durchbiegung des Trägerbalkens des Triangulationsaufbaus. Der maximale meßtechnisch relevante Fehler tritt bei der vergleichenden Betrachtung des waagrechten Trägerbalkens mit nach oben (Fall 1) und nach unten (Fall 2) gerichteten Kameras auf,

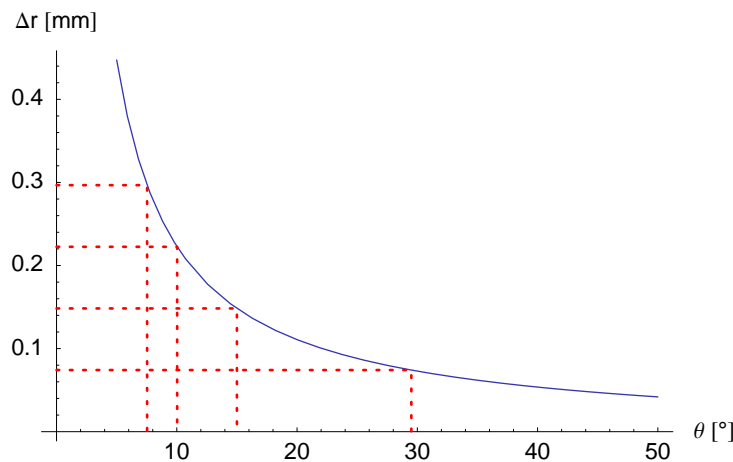


Abbildung 2.18: Auswirkung eines konstanten Winkelfehlers auf die Fehlerschranke der Entfernungsmessgenauigkeit in Abhängigkeit vom Triangulationswinkel.

da hier die größte relative Lageänderung der Kameras zueinander auftritt (absolute Lageänderungen beeinträchtigen nicht die Messgenauigkeit, da die genaue absolute Lage des Objekts wegen der Positionierungsgenauigkeit des Roboters ohnehin unbekannt ist). Abbildung 2.19 zeigt die betrachteten Konstellationen und beteiligten Größen $L = \frac{1}{2}b$ (Balkenlänge ab mittlerer Einspannung), s (Durchbiegung), ψ (Verdrehung) sowie die

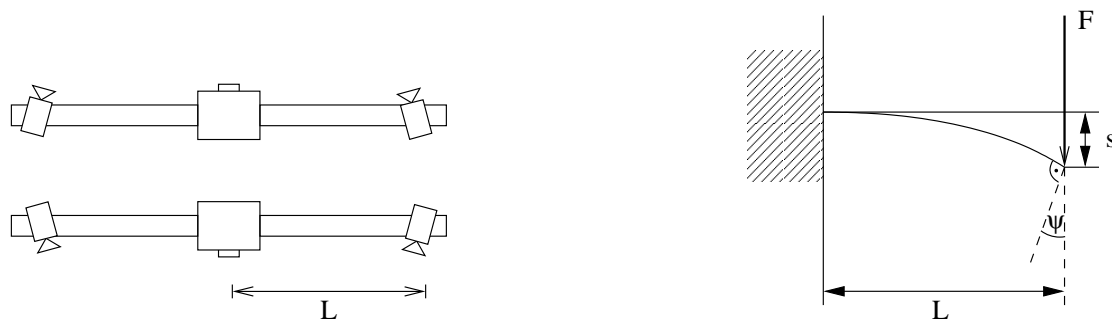


Abbildung 2.19: Biegung des Trägerbalkens.

einwirkende Kraft F (Gewicht der Kamera), zu der noch die gleichmäßig wirkende Kraft durch das Eigengewicht des Balkens hinzukommt.

Zur Berechnung des Winkels der Verdrehung wird die Durchbiegung des Balkens zunächst als Funktion $s : \mathbb{R} \rightarrow \mathbb{R}$ der Balkenlängsachse formuliert. Hierzu werden die Durchbiegungen durch Eigengewicht s_1 und Traglast s_2 mit

$$s_1 = \frac{F \cdot L^3}{8E \cdot I_a} \quad \text{und} \quad s_2 = \frac{F \cdot L^3}{3E \cdot I_a} \quad (2.24)$$

mit dem materialspezifischen Elastizitätsmodul E und dem von der Querschnittsgeometrie abhängigen Flächenträgheitsmoment I_a ausgedrückt [104, 111]. Setzt man nun an jedem Punkt x der Balkenlängsachse das Gewicht des von der Einspannung weiter entfernten Balkenabschnitts als zusätzliche Traglast an, so erhält man mit der Kraft c des Eigengewichts pro Längeneinheit die Durchbiegung

$$s(x) = \frac{c \cdot x^4}{8E \cdot I_a} + \frac{(F + c(L - x)) \cdot x^3}{3E \cdot I_a} \quad (2.25)$$

als Funktion des waagrechten Abstands zur Einspannung. Der Betrag der Durchbiegung selbst ist allerdings, anders als die Verdrehung, für die Genauigkeit der Entfernungsmessung irrelevant, da er in senkrechter Richtung gemessen wird und daher auf den symmetrischen Aufbau nur translatorisch wirkt, ohne die innere Geometrie zu verändern. Der Winkel ψ der Verdrehung kann aus der Steigung der Biegelinie mit

$$\tan \psi = s'(L) \quad (2.26)$$

abgeleitet werden. Als weitere mögliche Fehlerquelle kann die Verkürzung des waagrechten Abstands der Kamera von der Balkeneinspannung identifiziert werden. Über die Weglänge der Biegelinie kann durch Lösen der Gleichung

$$\int_0^{x_L} \sqrt{1 + s'(x)^2} \, dx = L \quad (2.27)$$

der tatsächliche waagrechte Abstand x_L berechnet werden [12].

Für handelsübliche Strangpreßprofile aus Aluminium können die relevanten Kennzahlen vom Hersteller bezogen und mit den verwendeten Abmaßen und dem Kameragewicht von 600 Gramm in die obenstehenden Gleichungen 2.26 und 2.27 eingesetzt werden. Dabei stellt sich zunächst heraus, daß die Verkürzung der Basislinie durch die Durchbiegung vernachlässigt werden kann, da sie bei allen betrachteten Ausführungen im Submikrometerbereich liegt und außerdem im betrachteten Fall neutralisiert wird (die Verkürzung ist für die Blickrichtung nach oben und nach unten identisch). Die berechneten Winkel der Verdrehung und die daraus mit Gleichung 2.23 resultierenden Fehlerschranken der Entfernungsmessung für einen Meßabstand von 1000 mm sind in Tabelle 2.1 aufgetragen.

Zur Diskussion der berechneten Fehlerschranken sei zunächst angemerkt, daß die Summierung der Einzelfehler $\Delta\alpha$ und $\Delta\beta$ in diesem Fall keine theoretische Abschätzung darstellt (wie die Betrachtung des Winkelfehlers aufgrund der Kamerakalibrierung), sondern auch geometrisch motiviert ist – der geschätzte Meßfehler tritt also tatsächlich auf. Dabei ist aufgrund der perspektivischen Kamerageometrie neben der translatorischen Auswirkung des Entfernungsmeßfehlers auch mit einer Verzerrung

Ausführung	Länge [mm]	Querschnitt	ϑ [Grad]	ψ [μRad]	Fehler Δr [μm]
Aluprofil ^a	261	30 × 30 mm ²	15.0°	43.03	651.08
Aluprofil	174	40 × 40 mm ²	10.0°	5.69	130.04
Aluprofil	261	40 × 40 mm ²	15.0°	12.94	195.84
Aluprofil	500 ^b	40 × 40 mm ²	18.8°	11.62	141.82
Aluprofil	261	50 × 50 mm ²	15.0°	5.90	89.27
CFK-Rohr ^c	261	Ø30 - 25 mm	15.0°	39.91	603.88
CFK-Rohr	261	Ø60 - 50 mm	15.0°	2.03	30.77

^aDie Ergebnisse beziehen sich auf feingerichtete Strangpreßprofile nach DIN 17615 von MayTec, für die der Hersteller jeweils Trägheitsmoment und Gewicht angibt [66]. Die Materialkonstante E wird für die verwendete Legierung Al Mg Si 0.5 F25 mit $E = 70 \text{ kN/mm}^2$ angegeben.

^bNur 326 mm werden als Basislinie genutzt; aufgrund der Einspannung beträgt die gebogene Länge auf beiden Seiten 110 mm. Der nicht genutzte Teil des Balkens muß der Traglast zugeschlagen werden.

^cMehrschicht-CFK-Verbund aus SIGRAFIL C® / Epoxidharz (60% Faseranteil), für den der Anbieter Gewicht und die Materialkonstante $E = 125 \text{ kN/mm}^2$ angibt [14]. Das Trägheitsmoment kann für Rohre mit Standardformeln [104] berechnet werden.

Tabelle 2.1: Beispielrechnungen für den Entfernungsmeßfehler durch Biegung des Trägerbalkens nach Gleichung 2.23. Die im Rahmen der vorliegenden Arbeit genutzte Variante ist hervorgehoben.

der erfaßten Geometrie zu rechnen. Dennoch darf angenommen werden, daß ein wesentlicher Anteil des Entfernungsmeßfehlers nur translatorische Auswirkungen auf die erfaßte Geometrie hat und daher bei der beabsichtigten Anwendung keine Rolle spielt.

Als wichtigstes Ergebnis der Fehlerberechnung kann der Zusammenhang der Biegeproblematik mit der Wahl des Triangulationswinkels ϑ festgehalten werden. Wie man an den Beispielen des 40 mm-Profils mit unterschiedlichen Triangulationswinkeln sieht, kann der Verlust an Meßgenauigkeit durch die Wahl eines kleineren Triangulationswinkels unter Umständen sogar kompensiert werden, weil sich der entsprechend kürzere Trägerbalken weniger stark verbiegt.

Generell erscheinen Aluminium-Profile als Trägermaterial für bewegte Triangulationsmeßsysteme problematisch. Weitere Beispielrechnungen zu Rohren aus kohlefaserverstärktem Kunststoff (CFK) zeigen, daß bei einem einfachen Balkenaufbau erst durch relativ hohe Trägerstärken eine signifikante Erhöhung der Stabilität erreicht wird. Für meßtechnische Anwendungen der robotergestützten 3D-Digitalisierung mit entsprechend hohen Genauigkeitsanforderungen erscheint daher die Konstruktion eines elastostatisch optimierten Trägersaufbaus sinnvoll, beispielsweise eines Fachwerks aus mehreren Balken. Im Rahmen der vorliegenden Arbeit wurde stattdessen der Fehler

durch ein 40 mm-Profil toleriert und durch einen größeren Triangulationswinkel sowie eine breitere Einspannung reduziert (Tabelle 2.1, Zeile 4).

2.3.3 Multifokale Systeme

Multifokale Meßsysteme können in zwei Unterklassen eingeteilt werden, von denen die erste auf mechanischen Komponenten aufsetzt, die zweite dagegen auf Kamerasystemen mit telezentrischen Optiken basiert. Mechanische Komponenten werden vor allem mit nativ punkt- oder zeilenweise erfassenden Sensoren kombiniert, während telezentrische Optiken mit allen kamera-basierten Verfahren kombiniert werden können. Im Folgenden sollen einige häufig anzutreffende Beispiele der Kombination unifokaler und triangulations-basierter Sensoren mit mechanischen Komponenten beispielhaft beschrieben werden.

Triangulations-basierte Laserlichtschnitt-Sensoren werden oftmals ohne mechanische Komponenten zur Variation der Laserebene angeboten. Der Aufbau reduziert sich damit auf eine Kamera mit geeignetem Filter und eine Zylinderlinse zur Aufweitung des Laserstrahls, so daß der Sensor nativ zu jedem Zeitpunkt eine Bildzeile parallel erfassen kann. Eine besonders kostengünstige Methode zur Erfassung zweidimensionaler Tiefenbilder besteht in der Ausnutzung einer häufig ohnehin vorhandenen Bewegung des zu vermessenden Objekts quer zur Bildzeile. Dieser Ansatz ist daher vor allem zur kontinuierlichen Profilerfassung an linear bewegten Stetigförderern (zum Beispiel Förderband) geeignet, bei denen die Förderbewegung die zweite Bildachse erzeugt. Die entstehende Bildgeometrie ist in der einen Achse durch eine perspektivische Projektion, in der anderen dagegen näherungsweise durch eine Parallelprojektion bestimmt.

Ein weiteres Beispiel der Kombination eines Lasertriangulations-Sensors mit mechanischen Komponenten zielt auf die möglichst vollständige allseitige Erfassung geometrisch einfacher Objekte ab: der Cyberware-Scanner bewegt einen zeilenweise erfassenden Laserlichtschnitt-Sensor auf einer kreisförmigen Trajektorie um das zu vermessende Objekt (insbesondere den menschlichen Kopf), so daß dieses in nur einem Tiefenbild abgesehen von Hinterschnitten vollständig erfaßt werden kann [25]. Hier ist die entstehende Bildgeometrie in der einen Achse durch eine perspektivische Projektion bestimmt, in der anderen Achse kann sie durch eine zylindrische Projektion mit Versatz zur Zylinderachse funktional beschrieben werden.

Zur Vermessung von Eisenbahn-Tunneln werden lichtlaufzeitbasierte Sensoren mit radialer Ablenkung durch rotierende Spiegel eingesetzt, die auf einer fahrenden Lokomotive montiert werden. Die native Geometrie des entstehenden Bildes ist daher in der einen Achse zylindrisch-spiralförmig, in der anderen Richtung durch die Trajektorie der befahrenen Eisenbahn-Schienen bestimmt [46]. Eine ähnliche Situation tritt auch bei

der Erfassung von Geländeprofilen mit Hilfe von Radar-Zeilensensoren vom Flugzeug aus [22] auf³. Dabei wird die Trajektorie des Sensors durch eine direkte Georeferenzierung mit differentielltem GPS und inertialen Sensoren gewonnen. Die auffälligste Eigenschaft derartiger Bilderfassungsmethoden besteht in der großen Genauigkeitsdifferenz, mit der die Kamerageometrie in den beiden Bildachsen bestimmt werden kann.

Die intrinsische Geometrie von Bildern, die von einem Zeilensensor und der Flugzeug-Trajektorie bestimmt wird, eignet sich auch zur Veranschaulichung einiger Einschränkungen der Verwendbarkeit von Bilddaten im Rahmen der vorliegenden Arbeit: Zum einen ist intuitiv einsichtig, daß die Zuordnung eines Punktes in Weltkoordinaten zum Bildraster mehrere Lösungen haben kann, wenn das Flugzeug eine Region mehrmals zum Beispiel auf einer Kreisbahn überfliegt. Zum anderen kann durch Änderung der Flugzeugneigung („Nicken“) lokal die Scan-Richtung gewechselt werden, so daß bereits vermessene Bereiche spiegelbildlich ein weiteres Mal im Bild repräsentiert werden. Mathematisch entsteht in der Abbildungsfunktion von 3D-Weltkoordinaten in den Rasterkoordinatenraum des Bildes an dieser Stelle eine Singularität.

2.3.4 Abstraktionsgrad und Schnittstelle

Die in den folgenden Kapiteln beschriebenen Verfahren zur automatisierten 3D-Digitalisierung setzen voraus, daß nicht allein das vom Sensor gelieferte Meßergebnis in Form einer Punktwolke, sondern auch die Geometrie seiner Erzeugung bekannt ist. Von den Sensoreigenschaften darf nur insoweit abstrahiert werden, als die rechnerische Prädiktion eines zusätzlichen Akquisitionsschrittes mit dem Sensor möglich bleibt.

Als sinnvolle Schnittstelle hat sich dabei eine funktionale Beschreibung der Sensorgeometrie herausgestellt, die die bijektive Abbildung zwischen kartesischen 3D-Koordinaten im Koordinatensystem des Sensors und dem zweidimensionalen Bildraster des Tiefenbildes mit dem gespeicherten Meßwert als (abhängiger) dritter Koordinate formalisiert. Durch die Einbeziehung des im Bildraster gespeicherten Meßwerts wird der funktionale Zusammenhang umkehrbar. Durch partielle Differenzierung ergeben sich weitere wichtige Eigenschaften des Bildrasters wie zum Beispiel die einem Bildpunkt zugeordnete Blickrichtung und die Winkelauflösung. Daher muß die Abbildungsfunktion im genutzten Bereich stetig differenzierbar und bijektiv sein. Eine genauere Beschreibung der genutzten Eigenschaften der Abbildungsfunktion erfolgt in Abschnitt 3.3.

³Im anglo-amerikanischen Sprachraum wird diese Konstellation auch als sog. „pushbroom“-Kameramodell bezeichnet. Derartige Systeme sind hier nur vollständigheitshalber erwähnt; sie sind für die vorliegende Arbeit nur von theoretischem Interesse.

Weiterhin müssen als meßtechnische Rahmenbedingungen die Grenzen des Meßbereichs in den lateralen Richtungen und in der Distanz-Meßrichtung bekannt sein. Für triangulations-basierte Sensoren erfordern die verwendeten Verfahren außerdem die Kenntnis aller beteiligten Beobachtungspunkte. Außerdem ist eine gerätespezifische Qualitätsmetrik sinnvoll, aber für das vorgeschlagene Verfahren nicht zwingend erforderlich, die aus dem Winkel zwischen Blickrichtung und Oberflächennormale und der Entfernung eine Maßzahl für die zu erwartenden Datenqualität liefert. Auch hier wird für triangulations-basierte Sensoren zusätzlich der Winkel zwischen Oberflächennormale und Blickrichtung des zweiten Beobachtungspunktes benötigt.

2.4 Automatisierungstechnik

Die Umsetzung algorithmisch bestimmter Aufnahmekonfigurationen erfordert ein geeignetes Handhabungsgerät, das Sensor und Objekt in eine vorzugebende relative Lage zueinander bringen kann. Die in Frage kommende Geräteklasse ist die der Industrieroboter, welche nach VDI-Richtlinie 2860 folgendermaßen definiert wird:

Industrieroboter sind universell einsetzbare Bewegungsautomaten mit mehreren Achsen, deren Bewegungen hinsichtlich Bewegungsfolge und Wegen beziehungsweise Winkeln frei programmierbar (d.h. ohne mechanischen Eingriff vorzugeben beziehungsweise änderbar) und gegebenenfalls sensorgeführt sind. Sie sind mit Greifern, Werkzeugen oder anderen Fertigungsmitteln ausrüstbar und können Handhabe- oder andere Fertigungsaufgaben ausführen [112].

Die auszuführende Aufgabe ist im vorliegenden Fall eine Inspektions- oder Vermessungsaufgabe, die, wenn es sich um unbekannte Einzelstücke handelt, mit „Losgröße 1“ durchgeführt werden muß. Einschränkend kommt hinzu, daß das Objekt erst durch den Vermessungsvorgang Schritt für Schritt bekannt wird. Daher genügt die Eigenschaft „frei programmierbar“ nicht (diese bezeichnet im Allgemeinen einen *off-line*-Vorgang mit Unterbrechung des normalen Programmablaufs), sondern die Bewegungen müssen während der Laufzeit des Verfahrens vorgegeben werden können.

In der industriellen Praxis werden heute fast ausschließlich stationäre Roboter eingesetzt, die einen bauartbedingt eingeschränkten Arbeitsbereich aufweisen. Diese eignen sich naturgemäß nur zur Digitalisierung von Werkstücken, deren Oberfläche (zumindest abschnittsweise) in diesen Arbeitsbereich gebracht werden kann. Zur Digitalisierung unbekannter Innenräume ist dagegen ein zumindest in zwei Dimensionen weitgehend unbeschränkter Arbeitsbereich notwendig. Im folgenden Abschnitt 2.4.1 werden daher zunächst Eigenschaften der in der Industrie üblichen stationären Roboter beschrieben, während in Abschnitt 2.4.2 die davon abweichenden Eigenschaften

mobiler Roboter-Lösungen beschrieben werden. Die bauartbedingte Einschränkung des Arbeitsbereichs bildet dabei das Unterscheidungskriterium.

2.4.1 Stationäre Roboter

Stationäre Industrieroboter sind üblicherweise als offene kinematische Kette realisiert, bei der jedes Armteil durch ein Dreh- oder Schubgelenk mit dem folgenden Armteil verbunden wird. Das dem ersten Gelenk vorangehende Armteil wird als Basis, das dem letzten Gelenk folgende als Effektor bezeichnet. Gelenke mit mehr als einem Freiheitsgrad wie zum Beispiel Kugelgelenke werden per Konvention durch einen sogenannten fiktiven Armteil ohne räumliche Ausdehnung verbunden, um eine einheitliche rechnerische Behandlung zu ermöglichen.

2.4.1.1 Beschreibung nach Denavit-Hartenberg

Durch die Winkel beziehungsweise Schublängen der Gelenke ist die Lage des Effektors relativ zur Basis eindeutig bestimmt. Zur kompakten Repräsentation der gesamten kinematischen Kette und ihres Zustands hat sich die *Denavit-Hartenberg-Konvention* [26] etabliert. Hierzu wird jedem Armteil einschließlich der Basis nach bestimmten Vorschriften ein Koordinatensystem zugeordnet, so daß die Orientierung⁴ des Effektors durch eine Kette von Koordinatensystemtransformationen bestimmt werden kann, die ihrerseits durch die Gelenkstellungen parametrisiert sind. Ziel der Denavit-Hartenberg-Konvention ist es, die relativen Orientierungen zwischen den Armteil-Koordinatensystemen mit nur vier Parametern (zwei Rotationswinkel und zwei Translationsparameter) angeben zu können, von denen jeweils einer von der repräsentierten Gelenkstellung abhängig ist.

Im folgenden wird die Konvention zur Festlegung der Armteil-Koordinatensysteme in Anlehnung an Weber [112] wiedergegeben. Für eine offene kinematische Kette mit n Gelenken wird jedem Armteil $i \leq n$ ein lokales Koordinatensystem K_i zugeordnet, das von den Basisvektoren x_i , y_i und z_i aufgespannt wird. Der Ursprung jedes Koordinatensystems K_i für $i < n$ liegt in der Gelenkachse $i + 1$; die weitere Orientierung wird folgendermaßen festgelegt:

Basiskoordinatensystem: Der Ursprung des Weltkoordinatensystem K_0 , auf das sich die Lage der Roboter-Basis bezieht, kann auf der Gelenkachse 1 frei gewählt werden. Die z_0 -Achse koinzidiert mit der Gelenkachse 1; die übrigen Basisvektoren können orthogonal zu z_0 unter Beachtung der Händigkeit frei gewählt werden.

⁴Hier und im Folgenden wird durch den Begriff „Orientierung“ die Rotation und Translation gegenüber einem Bezugskoordinatensystem (Freiheitsgrad 6) bezeichnet.

Armteil-Koordinatensysteme: Der Ursprung des Koordinatensystems K_i wird auf Gelenkachse $i + 1$ so gewählt, daß der Abstand zu Gelenkachse i minimal wird. Ist dieser Punkt nicht eindeutig definiert (weil die Gelenkachsen parallel verlaufen), so wird zunächst der Ursprung von K_{i+1} bestimmt, und der Ursprung von K_i auf Gelenkachse $i + 1$ so gewählt, daß der Abstand zum Ursprung von K_{i+1} minimal wird.

Der Basisvektor z_i koinzidiert mit der Gelenkachse $i + 1$, für den Basisvektor x_i wird eine Fallunterscheidung getroffen:

- Schneiden sich die Gelenkachsen $i - 1$ und i , so wird der Basisvektor x_i durch das Kreuzprodukt $z_{i-1} \times z_i$ aufgespannt.
- Andernfalls wird der Basisvektor x_i durch das Kreuzprodukt $z_i \times z_{i+1}$ gebildet und so orientiert, daß Achse $i + 1$ in positiver x_i -Richtung vom Ursprung von K_i liegt.

Der Basisvektor y_i ist bereits durch die Händigkeit des Koordinatensystems eindeutig bestimmt.

Effektor-Koordinatensystem: Die Lage des Effektor-Koordinatensystems K_n ist so zu wählen, daß die relative Orientierung zwischen K_{n-1} und K_n ebenfalls mit den vier im folgenden definierten Denavit-Hartenberg-Parametern beschrieben werden kann.

In der Praxis bedeutet dies, daß ein gegebenes Koordinatensystem des Werkzeugs (im Falle der vorliegenden Arbeit also das interne Koordinatensystem der Kamera, das durch das optische Zentrum, die Hauptachse und die Richtung der Bildachsen am Hauptachsenpunkt gebildet wird) im Allgemeinen nicht gemäß der Denavit-Hartenberg-Konvention als Effektorkoordinatensystem definiert werden kann. Daher wird zusätzlich zu K_n noch eine konstante affine Abbildung von K_n in das Werkzeugkoordinatensystem definiert.

Mit den so definierten Koordinatensystemen kann die relative Orientierung zweier Koordinatensysteme K_i und K_{i+1} durch die vier sogenannten *Denavit-Hartenberg-Parameter* angegeben werden. Diese sind der Drehwinkel ϑ_i um die Gelenkachse i , die Translation d_i , die dem Abstand der Ursprünge von K_{i-1} und K_i in Richtung z_{i-1} entspricht, die Translation a_i , die dem Abstand der Gelenkachsen $i - 1$ und i in Richtung x_i entspricht, sowie dem Drehwinkel α_i um x_i .

Damit kann die Abbildung vom Koordinatensystem K_{i-1} in das Koordinatensystem K_i durch die homogene Matrix

$$T^{(i)} = \begin{pmatrix} \cos \vartheta_i & -\sin \vartheta_i \cos \alpha_i & \sin \vartheta_i \sin \alpha_i & a_i \cos \vartheta_i \\ \sin \vartheta_i & \cos \vartheta_i \cos \alpha_i & -\cos \vartheta_i \sin \alpha_i & a_i \sin \vartheta_i \\ 0 & \sin \alpha_i & \cos \alpha_i & d_i \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (2.28)$$

angegeben werden; die Matrizen $T^{(i)}$ werden auch als *Denavit-Hartenberg-Matrizen* bezeichnet. Der variable Parameter ist bei einem Drehgelenk der Winkel ϑ_i , bei einem Schubgelenk dagegen die Schublänge d_i . Die gesamte Abbildung von Weltkoordinaten in Effektorkoordinaten läßt sich durch Verkettung der Abbildungen mit der homogenen Matrix

$$T_{Welt \rightarrow Effektor} = \prod_{i=1}^n T^{(i)} \quad (2.29)$$

darstellen. Üblicherweise werden aus dieser Darstellung die sogenannten Gelenkkoordinaten (auch als Roboterkoordinaten bezeichnet) abgeleitet, die in einem Vektor $\mathbf{q} \in \mathbb{R}^n$ repräsentiert sind, welcher den jeweils variablen Denavit-Hartenberg-Parameter q_i zu jedem Gelenk enthält.

2.4.1.2 Weitere Definitionen

Die einzelnen Achsen eines Roboters werden in Hauptachsen und Nebenachsen unterteilt. Die Hauptachsen sind zur Umsetzung signifikanter Positionsänderungen des Werkzeugs geeignet, während die Bewegungen der Nebenachsen nur kleine Positionsänderungen bewirken und primär die Rotation des Werkzeugs bestimmen. Schneiden sich die letzten drei Achsen der kinematischen Kette in einem Punkt, so wird dieser Teil der Kette als Zentralhand bezeichnet. Roboter mit mehr als sechs unabhängigen Achsen und einem entsprechenden Freiheitsgrad > 6 weisen eine sogenannte redundante Kinematik auf, die meist zur Vergrößerung des im Folgenden definierten Arbeitsbereichs genutzt wird.

Im Allgemeinen können die n Gelenke eines Roboters nur einen eingeschränkten Wertebereich von Drehwinkeln beziehungsweise Schublängen umsetzen; außerdem müssen gegebenenfalls Kollisionen mit anderen Armteilen oder der Umgebung vermieden werden, so daß weitere Teilbereiche des Gelenkkoordinatenraums nicht umgesetzt werden können. Die Gesamtheit aller erreichbaren Positionen des Effektors wird als Arbeitsraum des Roboters bezeichnet und ist in Abbildung 2.20 dargestellt. Die Teilmenge des Arbeitsraums, an der auch die Rotation des Effektors frei gewählt werden kann, wird als *dexterous workspace* bezeichnet.

2.4.1.3 Inverse Kinematik

Mit dem in Abschnitt 2.4.1.1 dargestellten Verfahren kann aus gegebenen Gelenkkoordinaten auf eindeutige Weise die Orientierung des Effektors und damit des Werkzeugs berechnet werden (direktes kinematisches Problem). Zur Umsetzung algorithmisch bestimmter Aufnahmekonfigurationen wird jedoch die umgekehrte Zuordnung

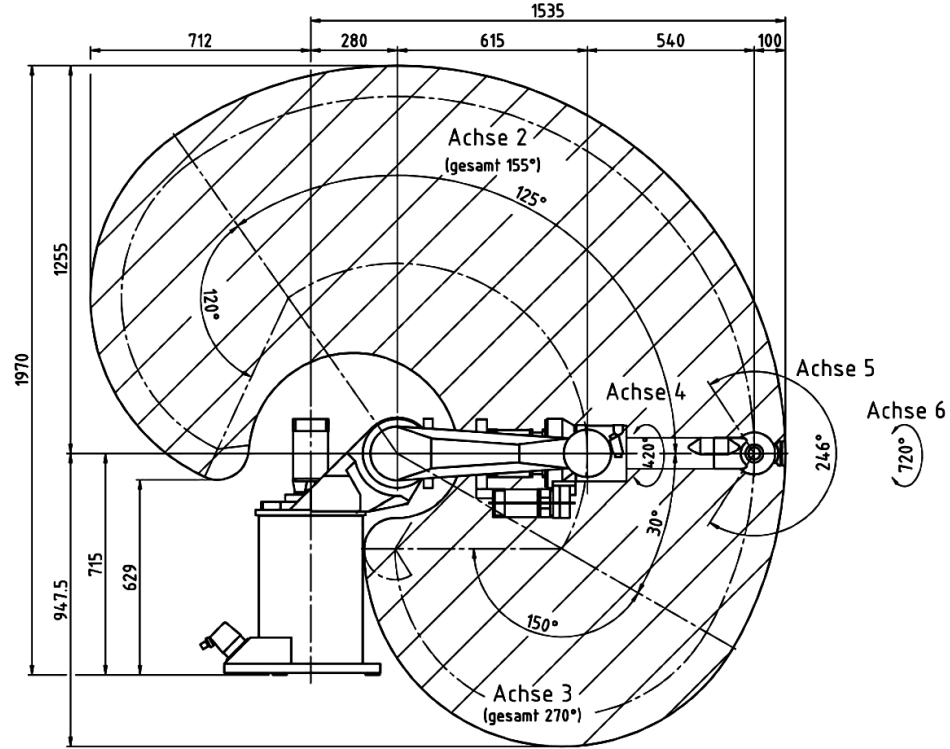


Abbildung 2.20: Arbeitsbereich eines sechssachigen Vertikalknickarmroboters (Reis RV16 [88, mit freundlicher Genehmigung]) mit Zentralhand. Der 3D-Arbeitsbereich ergibt sich durch Rotation der schraffierten Fläche um Achse 1, die senkrecht in der Mitte der Basis liegt, im Bereich $+150^\circ - -180^\circ$.

von Gelenkkordinaten zu einer gegebenen Orientierung des Werkzeugs benötigt (inverses kinematisches Problem) [45]. Die Lösung dieses Problems wird durch mögliche Mehrdeutigkeiten und Singularitäten erschwert.

Zur formalen Beschreibung dieser Problematik wird zunächst die Orientierung zwischen Weltkoordinaten und Effektorkoordinaten durch einen sechselementigen Vektor $\mathbf{x} \in \mathbb{R}^6$ mit $\mathbf{x} = (\varphi_x, \varphi_y, \varphi_z, t_x, t_y, t_z)^T$ beschrieben, welcher elementar in eine homogene Matrix mit

$$S_{\text{Welt} \rightarrow \text{Effektor}}(\mathbf{x}) := \begin{pmatrix} \cos \varphi_z \cos \varphi_y & \cos \varphi_z \sin \varphi_y \sin \varphi_x - \sin \varphi_z \cos \varphi_x & \cos \varphi_z \sin \varphi_y \cos \varphi_x + \sin \varphi_z \sin \varphi_x & t_x \\ \sin \varphi_z \cos \varphi_y & \sin \varphi_z \sin \varphi_y \sin \varphi_x - \cos \varphi_z \cos \varphi_x & \sin \varphi_z \sin \varphi_y \cos \varphi_x - \cos \varphi_z \sin \varphi_x & t_y \\ -\sin \varphi_y & \cos \varphi_y \sin \varphi_x & \cos \varphi_y \cos \varphi_x & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (2.30)$$

überführt werden kann. Analog kann auch zu jedem Gelenk i mit Gleichung 2.28 eine

Funktion zur Berechnung der Denavit-Hartenberg-Matrix aus dem variablen Denavit-Hartenberg-Parameter q_i des Gelenks i mit $T_{DH}(q_i) := T^{(i)}$ angegeben werden, so daß die vollständige Abbildung von Gelenkkoordinaten in eine homogene Matrix mit

$$T_{Welt \rightarrow Effektor}(\mathbf{q}) := \prod_{i=1}^n T_{DH}(q_i) \quad (2.31)$$

angegeben werden kann. Das direkte kinematische Problem kann nun durch die Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}^6$ der Abbildung von Gelenkkoordinaten in die Parameter der relativen Orientierung des Effektors zur Basis mit

$$f(\mathbf{q}) := \{\mathbf{x} \mid T_{Welt \rightarrow Effektor}(\mathbf{q}) = S_{Welt \rightarrow Effektor}(\mathbf{x})\} \quad (2.32)$$

beschrieben werden. Die analytische Berechenbarkeit der Funktion f wurde im Wesentlichen durch die Gleichungen 2.28, 2.29 und 2.31 gezeigt⁵; die Lösung ergibt eine eindeutige Matrix $S_{Welt \rightarrow Effektor}(\mathbf{x})$.

Das inverse kinematische Problem kann nun durch die Umkehrfunktion $f^{-1} : \mathbb{R}^6 \rightarrow \mathbb{R}^n$ beschrieben werden, die die Abbildung von den Parametern der relativen Orientierung des Effektors zur Basis in Gelenkkoordinaten analog mit

$$f^{-1}(\mathbf{x}) := \{\mathbf{q} \mid T_{Welt \rightarrow Effektor}(\mathbf{q}) = S_{Welt \rightarrow Effektor}(\mathbf{x})\} \quad (2.33)$$

beschreibt. Eine Singularität bezüglich der Abbildung f tritt im Gelenkkoordinatenraum an denjenigen Stellen auf, für die die Jacobi-Matrix J_f von f

$$J_f = \frac{\partial f(\mathbf{q})}{\partial \mathbf{q}} \quad (2.34)$$

singulär ist. Dies ist beispielsweise der Fall, wenn zwei Gelenkachsen aufgrund der Stellung der dazwischenliegenden Gelenke koinzidieren, so daß ihre Gelenkwinkel frei wählbar sind.

In vielen Fällen hat $f^{-1}(\mathbf{x})$ mehrere Lösungen. So kann beispielsweise für Drehgelenke i mit einem Arbeitsbereich von mehr als 2π unter Umständen eine alternative Lösung \mathbf{q}' mit $|q_i - q'_i| = 2\pi \cdot N$, $N \in \mathbb{N}$ gefunden werden.

Sofern die nicht-redundante Kinematik eines Roboters eine Zentralhand aufweist, kann die Lösung von $f^{-1}(\mathbf{x})$ analytisch bestimmt werden [78]. Mehrdeutigkeiten müssen dabei nach festzulegenden Regeln aufgelöst werden, die beispielsweise den Abstand zum vorangehenden Punkt im Gelenkkoordinatenraum nach einer geeigneten

⁵Die bekannte Methode der Bestimmung der Elemente von \mathbf{x} aus einer homogenen Matrix wird vorausgesetzt. Diese sollte zur Vermeidung von numerisch instabilen Lösungen eine geeignete Auflösung der Mehrdeutigkeiten für φ_x und φ_z im Fall $|\sin \varphi_y| = 1$ beinhalten.

Metrik minimieren. Dabei dürfen natürlich nur solche Lösungen in Betracht gezogen werden, die innerhalb des Arbeitsbereichs der jeweiligen Gelenke liegen und die innerhalb des Gesamt-Arbeitsbereichs des Roboters auch umgesetzt werden können.

Zur Lösung des inversen kinematischen Problems für redundante Kinematiken und für kinematische Ketten ohne Zentralhand müssen numerische Optimierungsverfahren eingesetzt werden. Wegen des damit verbundenen Rechenaufwands und der Schwierigkeit, die Echtzeit-Bedingung einzuhalten, sind Gelenkarmroboter ohne Zentralhand in der Praxis nur selten anzutreffen; zusätzliche Achsen müssen meist explizit getrennt angesteuert werden, so daß die Problematik redundanter Kinematiken rechnerisch vermieden wird.

2.4.1.4 Praxis der Ansteuerung

Die Umsetzung zur Laufzeit festgelegter Orientierungen des Effektors wird von der Steuerungssoftware marktgängiger Robotersysteme oftmals nur eingeschränkt unterstützt. Die Ursache dieser Einschränkung liegt in der Schwierigkeit, bei der Lösung des inversen kinematischen Problems zu konsistenten Lösungen zu kommen. Insbesondere wenn Mehrdeutigkeiten in Abhängigkeit vom Vorzustand aufgelöst werden, kann das Verhalten des Roboters zur Laufzeit nur noch schwer vorhergesagt und überprüft werden. Methoden der Pfadplanung, die dieses Problem im Prinzip lösen könnten, setzen eine weitgehende Kenntnis der Geometrie möglicher Hindernisse im Arbeitsraum voraus und sind daher für den industriellen Einsatz nicht praktikabel.

Das im Rahmen der vorliegenden Arbeit verwendete Robotersystem von Reis ermöglicht nicht die Definition neuer Effektororientierungen zur Laufzeit, sondern erlaubt lediglich die Anwendung eines gemischt kartesisch und gelenkkoordinatenweise angegebenen Korrekturvektors $\delta_{\mathbf{x}} \in \mathbb{R}^n$ mit

$$\delta_{\mathbf{x}} := (\varphi_x, \varphi_y, \varphi_z, t_x, t_y, t_z, x_7, \dots, x_n)^T, \quad (2.35)$$

der eine *off-line* definierte Effektororientierung modifiziert (nur der sechssachsige Gelenkarmteil des Roboters kann kartesisch korrigiert werden, die weiteren Achsen müssen mit direkten Gelenkwinkel- beziehungsweise Schublängen-Werten korrigiert werden). Der Korrekturvektor $\delta_{\mathbf{x}}$ wird typischerweise zur Laufzeit aus Sensorwerten gebildet, um beispielsweise den Programmablauf an die sensorisch erfaßte Lage eines nur ungenau positionierten Werkstücks anzupassen. Im Rahmen der vorliegenden Arbeit wird die XML-basierte TCP/IP-Schnittstelle der Robotersteuerung genutzt, um die Elemente des Korrekturvektors zur Laufzeit zu modifizieren.

Die Beträge der Elemente des Korrekturvektors sind abhängig von der Gelenkkoordinaten-Position des Roboters eingeschränkt; werden Korrekturen spezifiziert, die

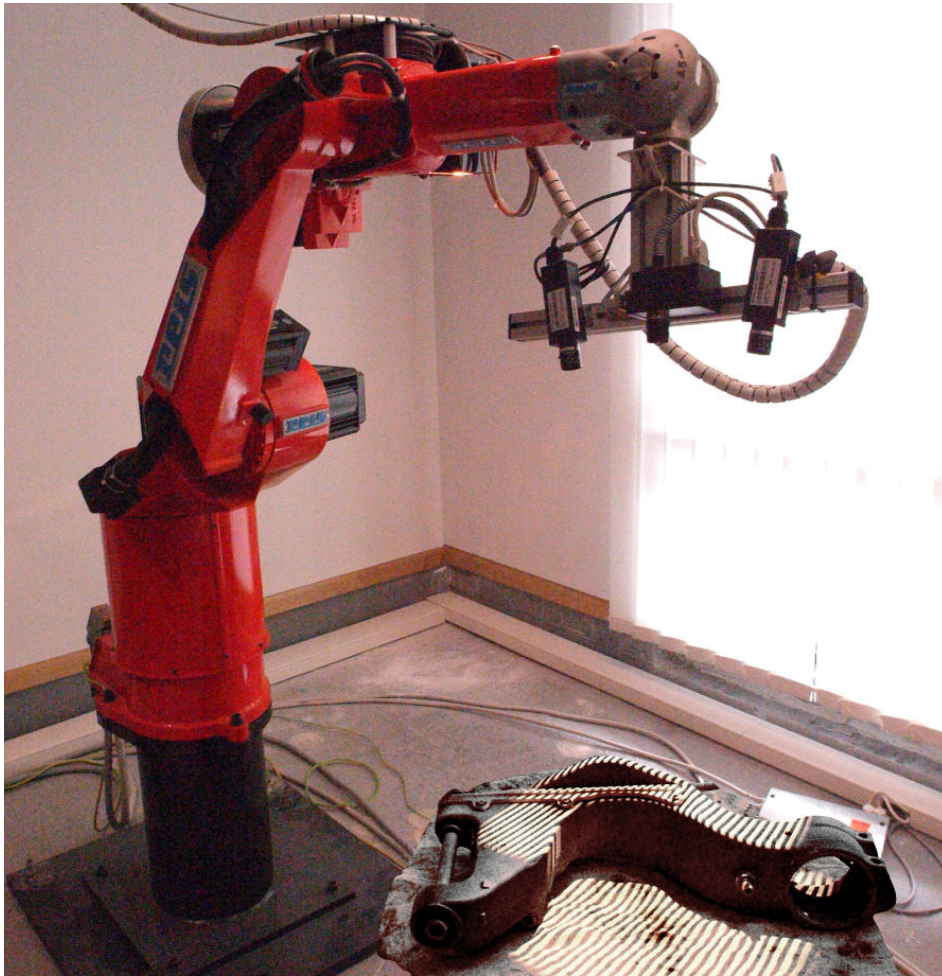


Abbildung 2.21: Gelenkarmroboter mit Streifenlichtsensor und vermessenem Objekt auf Drehteller.

bezüglich der definierten Effektororientierung eine signifikante Änderung der Gelenkstellungen bewirken, eine andere Auflösung von Mehrdeutigkeiten vorsehen oder den Durchgang von Bereichen des Gelenkkoordinatenraums erfordern, in denen die Jacobi-Matrix schwach konditioniert oder singulär ist, so verursacht ihre Anwendung eine Fehlermeldung von Seiten der Robotersteuerung. Dies ist mit Blick auf den beabsichtigten automatischen Ablauf der 3D-Digitalisierung höchst unerwünscht, da Fehlermeldungen der Robotersteuerung nicht softwareseitig abgefangen werden können, sondern zu einer Unterbrechung des Programmablaufs führen.

Eine softwareseitige Prüfung, ob ein gegebener Korrekturvektor bezogen auf eine vordefinierte Effektororientierung zu einem Fehler führt oder nicht, ist ebenfalls nicht möglich. Der Algorithmus zur Lösung des inversen kinematischen Problems wird von

den Herstellern der Robotersteuerung in der Regel weder zugänglich gemacht noch in seinen Eigenschaften dokumentiert. Aufgrund unterschiedlicher Genauigkeiten der verwendeten numerischen Prozessoren kann das genaue Verhalten der Robotersteuerung meist auch nicht zuverlässig vorab simuliert werden [112]. Der algorithmische Lösungsansatz wurde daher als undurchführbar verworfen.

Als Lösung wird dagegen ein Feld von vordefinierten Effektororientierungen vorgeschlagen, in dem die relativen Abstände so klein sind, daß jede relevante Orientierung durch Anwendung eines Korrekturvektors auf die jeweils nächstliegende vordefinierte Effektororientierung erreicht werden kann. Gleichzeitig kann jeder vordefinierten Effektororientierung eine durch den Korrekturvektor erreichbare Umgebung zugeordnet werden, so daß die Umsetzbarkeit einer betrachteten Aufnahmeconfiguration softwareseitig auf einfache Weise geprüft werden kann.

Im Rahmen der vorliegenden Arbeit wurde die erreichbare Umgebung durch ein sehr einfaches Verfahren angenähert: ein experimentell bestimmter Vektor der Form

$$\delta_{\mathbf{x}}^{\max} := (\varphi_x, \varphi_y, \varphi_z, t_x, t_y, t_z)^T, \quad (2.36)$$

limitiert elementweise die erlaubte Abweichung von der nächstliegenden vordefinierten Effektorposition. Bei hinreichend dichten Effektorpositionen läßt sich so auf einfache Weise ein Kriterium aufbauen, welches Fehlermeldungen der Robotersteuerung aufgrund ungültiger Konfigurationen wirksam verhindert. Gleichzeitig kann auf diese Weise ein Kriterium zur Evaluation von Nebenbedingungen der Aufnahmeplanung bereitgestellt werden, welches nicht erreichbare Roboterpositionen bereits bei der Planung ausschließt.

2.4.2 Mobile Roboter

Mobile Roboter unterscheiden sich von stationären Robotern im Wesentlichen dadurch, daß das Basiselement ihres kinematischen Aufbaus nicht fest mit dem Weltkoordinatensystem verbunden, sondern auf einer mobilen Plattform montiert ist. Diese ist anwendungsspezifisch fahrend [21, 4], schwimmend oder fliegend [28], in seltenen Fällen auch auf Beinen laufend [3, 50] oder raupenartig kriechend [77] ausgeführt. Der wichtigste technische Unterschied zu stationären Systemen besteht darin, daß die aktuelle Orientierung der Plattform im Weltkoordinatensystem meist nur mit vergleichsweise geringer Genauigkeit bekannt ist. Der Arbeitsbereich ist bezüglich der Effektorposition in den beiden horizontalen Dimensionen typischerweise durch den mitgeführten Energievorrat und die geometrischen Eigenschaften der Umgebung, aber nicht durch den Roboter selbst eingeschränkt. In der vertikalen Dimension ist der Arbeitsbereich der mobilen Plattform im Regelfall eingeschränkt (Flughöhe, Tauchtiefe,



Abbildung 2.22: Mobiles Roboterfahrzeug, das die 3D-Lichtlaufzeitsensorik auf einer höhenverstellbaren Plattform (Schubgelenk) trägt.

feste Höhe über dem Untergrund); der kinematische Aufbau auf der Plattform unterliegt denselben Einschränkungen wie im Falle statischer Roboter.

Für mobile Roboter, die, wie im Falle der vorliegenden Arbeit, in initial unbekannten Umgebungen eingesetzt werden sollen, ist zur Umsetzung algorithmisch bestimmter Aufnahmekonfigurationen nicht nur das inverse kinematische Problem zu lösen, sondern zusätzlich das Problem der Pfadplanung in einer nur teilweise bekannten Umgebung, so daß die beabsichtigte Aufnahmeposition kollisionsfrei angefahren werden kann. Eine funktionierende Implementierung zur Lösung dieses Problems konnte im Rahmen der vorliegenden Arbeit genutzt werden, ohne daß eine Auseinandersetzung mit den Details oder spezifischen Schwierigkeiten notwendig geworden wäre. Auf eine detailliertere Beschreibung der Problematik wird daher an dieser Stelle verzichtet.

Es sei noch angemerkt, daß die erwähnte relativ ungenaue Positionsbestimmung der mobilen Plattform, die bei fahrenden Ausführungen meist durch Odometrie geschätzt werden muß, im Zusammenspiel mit dem transportierten Werkzeug zur flächenhaften optischen Oberflächenvermessung durch Registrierung der erfaßten Oberflächenbereiche erheblich verbessert werden kann [4]. Für schwimmende und fliegen-

de Plattformen ist ein derartiger Bezug auf externe Referenzen wegen der Dynamik des jeweiligen Mediums (Wind, Strömung) zur Navigation ohnehin unvermeidlich.

2.4.3 Abstraktionsgrad und Schnittstelle

Im Rahmen der vorliegenden Arbeit wurden zur Umsetzung der zur Laufzeit spezifizierten Aufnahmekonfigurationen Schnittstellen implementiert beziehungsweise genutzt, die eine direkte Vorgabe der benötigten Kameraorientierung in sechs Freiheitsgraden erlauben. Zur Beschleunigung der vorgeschlagenen Verfahren ist es darüberhinaus hilfreich, die Erreichbarkeit von Aufnahmekonfigurationen durch das verwendete Robotersystem mit möglichst geringem Aufwand vorab prüfen zu können. Hierfür nutzen die vorgeschlagenen Verfahren weitere Schnittstellen, die Information darüber zur Verfügung stellen, ob eine gegebene Aufnahmekonfiguration durch den Roboter erreichbar ist oder nicht.

Die Geometrie des verwendeten Roboters kann daher für die vorgeschlagenen Verfahren ebenso transparent bleiben wie die technischen Details der Ansteuerung, so daß das verwendete Robotersystem auf einfache Weise ausgetauscht werden kann.

2.5 Zusammenfassung eigener Ergebnisse

In diesem Kapitel wurde ein Überblick über die Gerätetechnik aus den Bereichen der optischen Meßtechnik und der Automatisierungstechnik gegeben, die sich zur Umsetzung der in der vorliegenden Arbeit vorgestellten Verfahren eignet. Hierzu wurde eine eigene Klassifikation flächenhaft messender 3D-Sensoren vorgeschlagen, die auf der jeweiligen Aufnahmeposition zu den erfaßten Datenpunkten basiert. Dieses Klassifikationskriterium ist durch die Einbeziehung der jeweiligen Aufnahmepositionen in die nachfolgend vorgeschlagenen Verfahren zur Aufnahmeplanung motiviert. Geeignete Schnittstellen zur verwendeten Sensorik wurden vorgeschlagen, so daß lediglich die Klassifizierung des Sensors spezifisch berücksichtigt werden muß, ansonsten aber vom konkret eingesetzten Gerät abstrahiert werden kann.

Triangulations-basierte Systeme bilden einen Schwerpunkt des Kapitels, der durch die besondere Situation der Messung mit mehr als einem involvierten Beobachtungspunkt motiviert ist. Hierzu wurden eigene Vorschläge zur numerischen Vereinfachung der Kalibrierung beschrieben sowie eine detaillierte Analyse der Meßfehler in Abhängigkeit von mechanischen Einflüssen durchgeführt, wie sie bei der Bewegung des Sensors durch einen Roboter auftreten.

Geeignete Roboter können im Vergleich zur Sensorik auf sehr viel einfachere Weise angebunden werden, da die in der vorliegenden Arbeit vorgeschlagenen Verfahren zur

Aufnahmeplanung eine Ausnutzung gerätespezifischer Konstruktionsprinzipien nicht vorsehen. Stattdessen stehen insbesondere bei Industrierobotern die Probleme der technischen Umsetzung der Ansteuerung im Vordergrund, da diese durch Erfordernisse der Arbeitssicherheit erschwert wird. Hierzu wurde ein Verfahren vorgeschlagen, das durch Bezugnahme auf bekannte Roboter-Posen eine prädizierbare und damit sichere Bewegung bei der Umsetzung der zur Laufzeit bestimmten Aufnahmekonfigurationen erlaubt.

KAPITEL 3

Objektrepräsentation

3.1 Einleitung

Das zu rekonstruierende Objekt ist während des inkrementellen Rekonstruktionsvorgangs teil- und näherungsweise durch die Meßdaten bekannt. Diese Meßdaten müssen zunächst in geeigneter Weise aufbereitet werden, damit die Planung weiterer Aufnahmen effizient durchgeführt werden kann. An die Repräsentation des teil- und näherungsweise vermessenen Objekts können drei wichtige Forderungen gestellt werden:

- Neben dem bereits erfaßten Teil der Oberfläche muß auch der noch unvermessene Teil des Objekts repräsentiert werden.
- Die Qualität der Vermessung muß zu jedem Punkt der Objektoberfläche repräsentiert werden.
- Die inkrementelle Integration zusätzlicher Aufnahmen muß mit geringem Aufwand möglich sein.

Ein zentraler Aspekt der Auswahl einer geeigneten Repräsentationsform ist die effiziente Sichtbarkeitsbestimmung, da diese bei der Evaluierung eines Parametersatzes zur nächsten Aufnahme den größten Anteil des Rechenaufwandes benötigt. Für die Sichtbarkeitsbestimmung steht heute leistungsstarke und kostengünstige Spezialhardware (hardware-beschleunigte 3D-Graphikkarten) zur Verfügung, die durch die Objektrepräsentation unterstützt werden sollte. Daher kommen im Rahmen der vorliegenden Arbeit nur diskret-geometrische Oberflächenmodelle in Frage, die durch computergraphische Lichtquellen schattiert und durch Definition von Materialeigenschaften (insbesondere Farbe) an jedem Punkt der Oberfläche in ihrer Darstellung differenziert werden können.

Die Einschränkung auf diskret-geometrische Oberflächenmodelle bietet auch dann Vorteile durch den reduzierten Rechenaufwand, wenn aufgrund der Sensorgeometrie keine durch Hardwareimplementierungen darstellbaren Blickfelder vorliegen, wie dies bei multifokalen Sensoren der Fall ist. Derartige Sensoren werden im Rahmen der vorliegenden Arbeit zwar nicht betrachtet, können aber durch Implementierung speziell an die Sensorgeometrie angepaßter Darstellungsverfahren ebenfalls unterstützt werden.

In den folgenden Abschnitten dieses Kapitels wird zunächst eine mathematische Definition der teilvermessenen Objektoberfläche hergeleitet, die sowohl auf die interne (Sensor innerhalb des zu rekonstruierenden Objekts) wie auch auf die externe (Sensor außerhalb des Objekts) Rekonstruktion angewendet werden kann. Anschließend werden zwei Verfahren zu ihrer inkrementellen Konstruktion sowie Verfahren zur Integration unvermessener Objektteile und qualitativer Maßzahlen beschrieben. Den Abschluß des Kapitels bilden Vorschläge zur Repräsentation des unbekannten Objekts zu Beginn des Verfahrens sowie eine kurze vergleichende Zusammenfassung der 3D-Rekonstruktion auf Basis der automatisch akquirierten Einzelaufnahmen.

3.2 Stand der Technik

Die Erzeugung volumetrischer Objektbeschreibungen als Approximation der realen Geometrie findet sich in der Literatur zuerst im Zusammenhang mit Intensitätsbildern bei der Anwendung des Silhouettenverfahrens. Hierzu wird in mehreren Bildern das Objekt vom Hintergrund segmentiert. Zusammen mit den bekannten Aufnahmepositionen kann aus den segmentierten Bildern durch Vereinigung der eingesehenen Volumina eine Abschätzung der Objektgeometrie gewonnen werden. Martin und Aggarwal [63] thematisieren dabei die Anforderung der effizienten Modifikation durch zusätzliche Aufnahmen bei gleichzeitigem Erhalt der Detailinformation. Laurentini [56] unterscheidet in der Ergebnisgeometrie Bereiche, die bekanntermaßen auf der Objektoberfläche liegen, von solchen, die hilfsweise zur Ergänzung der geschlossenen Oberfläche hinzugefügt wurden.

Im Zusammenhang mit der Geometrie-Rekonstruktion aus mehreren Tiefenbildern sind die Ansätze des sogenannten *space carvings* zu nennen. Curless und Levoy [24] entfernen durch „Schnitzen“ zwischen Beobachtungspunkt und Meßpunkt Volumenanteile, die bekanntermaßen leer sind. Das verbleibende Volumen wird klassifiziert in „nicht eingesehen“ und „oberflächennah“. Der oberflächennahe Anteil wird bei der Isoflächenextraktion unter Berücksichtigung redundanter Meßdaten berücksichtigt, während die Grenzfläche zwischen „leer“ und „nicht eingesehen“ mit sogenannten „Lückenfüller“-Oberflächenelementen repräsentiert wird, so daß im Ergebnis eine wasserdichte Oberfläche konstruiert wird.

Einen ähnlichen Ansatz verfolgt Neugebauer [71], vermeidet aber die rechenintensive Isoflächenextraktion bei der Oberflächendefinition. Stattdessen schlägt er eine vorzeichenbehaftete Distanzfunktion vor, die effizient berechnet werden kann, weil die Tiefenbilddaten bereits durch einen Vorverarbeitungsschritt miteinander abgeglichen werden. Die Nulldurchgänge dieser Distanzfunktion werden trianguliert.

Pito [80] schlägt als Repräsentationsform des eingesehenen Volumens ein Voxelgrid vor; dieser Ansatz wird auch in vielen der früheren Arbeiten zur Aufnahmeplanung verfolgt: Banta et al. [6] schlagen ebenfalls ein Voxelgrid vor, Conolly [18] verwendet eine Octree-Datenstruktur.

In der vorliegenden Arbeit wird die vermessene und die unvermessene Objektoberfläche ähnlich dem Ansatz von Neugebauer durch Triangulation einer vorzeichenbehafteten Distanzfunktion erzeugt. Da für die Aufnahmeplanung aber eine niedrigere Datenqualität ausreichend ist, wird auf die in [72] vorgeschlagene aufwändige Vorverarbeitung der Tiefenbilddaten verzichtet.

Die wesentliche Neuerung des in der vorliegenden Arbeit vorgeschlagenen Verfahrens zur Konstruktion der Grenzfläche des eingesehenen Volumens besteht in ihrer inkrementellen Konstruktion. Im Gegensatz zu Pito [80] wird nicht das Voxelgrid bei der Integration zusätzlicher Aufnahmen verändert, sondern direkt das Dreiecksnetz. Dieser Ansatz ermöglicht den Einsatz relativ aufwändiger Verfahren zur Vereinfachung des Dreiecksnetzes (*mesh reduction*), da diese auf jedem neu hinzugefügten Abschnitt nur einmalig ausgeführt werden müssen, während bei einem inkrementell veränderten Voxelgrid eine vollständige Neuerzeugung des Dreiecksnetzes notwendig ist, da dessen ursprüngliche Struktur durch Reduktion zerstört wird.

3.3 Definition der teilvermessenen Objektoberfläche

Die Oberfläche des teilvermessenen Objekts ist durch die akquirierten Tiefenbilder gegeben, welche als teilweise definierte explizite Oberflächen $r(u, v)$ betrachtet werden. Die zunächst meßpunktweise an den ganzzahligen Pixelpositionen definierte Funktion r kann durch Interpolation stetig ergänzt werden; außerhalb der Bildränder sowie in Bereichen, in denen kein Meßpunkt vorliegt, ist sie undefiniert (Abbildung 3.1). Der Parameterraum der Datenpunkte des Tiefenbildes $(u, v, r(u, v))^T$ ist dabei durch das Bildraster und die interne Repräsentation der Entfernungswerte zu jedem Pixel gegeben und hängt daher in seiner geometrischen Bedeutung von den Sensoreigenschaften ab.

Dieser Zusammenhang wird durch die Projektionsfunktion $h : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ modelliert, welche einem Punkt im kartesischen Koordinatensystem des Sensors $\mathbf{x} \in \mathbb{R}^3$ die entsprechende Pixelposition im Bildraster sowie den dort vorliegenden Pixelwert

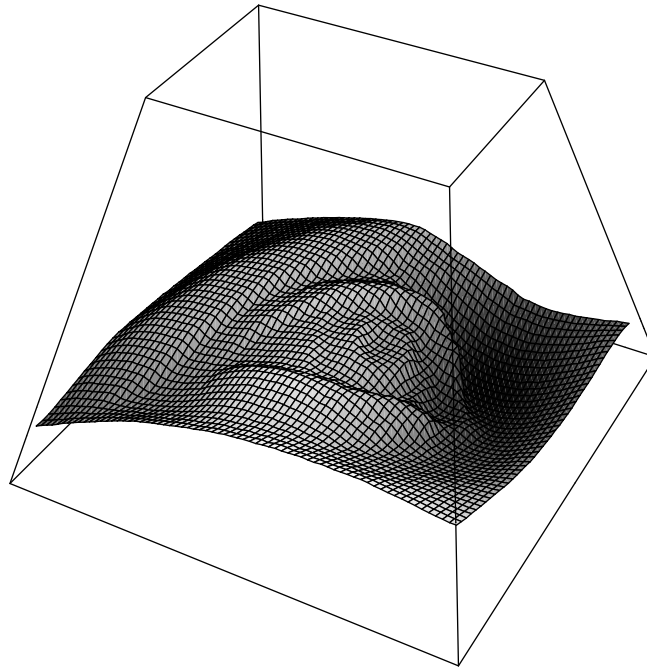


Abbildung 3.1: Die reliefartig erfaßte Oberfläche im Raster eines Tiefenbildes, durch Interpolation zwischen den Meßpunkten stetig ergänzt. Der Pyramidenstumpf und die Gitterlinien veranschaulichen die typische Geometrie der Parametrisierung eines perspektivischen Kamerabildes.

zuordnet. Die Funktion h hängt von den intrinsischen Parametern des Sensors ab, die typischerweise vorab bei der Kalibrierung bestimmt werden. Eine zentrale Eigenschaft der Funktion h ist ihre Umkehrbarkeit, so daß ein Meßpunkt des Tiefenbildes $(u, v, r(u, v))^T$ in kartesische Koordinaten überführt werden kann. Durch lokale Linearisierung der Funktion h mit Berechnung ihrer Jacobi-Matrix J_h durch

$$J_h = \frac{\partial h(x)}{\partial x} \quad (3.1)$$

können auch Vektoren, wie zum Beispiel die im Bildraaster berechneten Tangentialvektoren oder die Kamerablickrichtung, in kartesische Koordinaten überführt werden. Zur Vereinfachung der Notation wird im Folgenden für alle Bilder eine identische Funktion h angenommen, obwohl dies keine notwendige Nebenbedingung des Verfahrens ist.

Die zugrundeliegende Speicherung der Meßpunkte als Skalare in ihrem nativen Bildraaster hat gegenüber der expliziten Speicherung der 3D-Koordinaten der Meßpunkte einige Vorteile, deren offensichtlichster der geringe Speicherplatzbedarf ist: anstelle von drei Fließkommawerten genügt für die Genauigkeit und Meßbereiche ty-

pischer Sensoren ein einziger 16-bit Ganzzahlwert pro Meßpunkt¹. Im Falle großer Datensätze wiegt der Vorteil des schnelleren Zugriffs (die Daten können vollständig im Hauptspeicher gehalten werden) den höheren Rechenaufwand bei weitem auf. Im Folgenden wird darüber hinaus deutlich, wie die Aufnahmeplanung die Kenntnis der Sensorgeometrie durch die Funktion h ausnutzen kann. Neugebauer [72] zeigt, daß diese Vorgehensweise auch ohne explizite Kenntnis der Details der Sensorgeometrie für beliebige Sensortypen möglich ist, indem die Sensorgeometrie im Bildraster pixelweise gespeichert wird. Ein ähnlicher Ansatz wird auch von Grossberg und Nayar [36, 37] vorgeschlagen.

Die extrinsischen Parameter der Einzelbilder werden für jedes Bild i durch die Rotation und Translation (R_i, t_i) des Sensors relativ zum gemeinsamen Weltkoordinatensystem repräsentiert. Diese Parameter sind im Falle einer automatisierten Vermessung meist näherungsweise durch die bekannte Orientierung der Positioniereinrichtung bekannt und können durch Registrierung aus den Meßpunkten oder mit Hilfe eingemessener Marken vollautomatisch mit hoher Genauigkeit bestimmt werden. Für einen 3D-Punkt $\mathbf{x} \in \mathbb{R}^3$ in Weltkoordinaten ergibt sich damit der korrespondierende Datenpunkt $(u, v, w)^T$ im Bildraster durch

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = h(R_i \cdot \mathbf{x} + t_i) \quad (3.2)$$

Aus den vorliegenden registrierten Tiefenbildern r_i wird nun mit Hilfe der so definierten Abbildungen eine vorzeichenbehaftete Funktion $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ gebildet, die eine binäre Aufteilung des Raumes in den Bereich außerhalb des Objekts und den Bereich, der potentiell innerhalb des Objekts liegt, erlaubt. Dieser Aufteilung liegt die Idee zugrunde, daß Punkte zwischen der vermessenen Oberfläche und der Position des beobachtenden Sensors wegen der Sichtbarkeit der Oberfläche nicht innerhalb des Objekts liegen können, da diese andernfalls eine Verdeckung bilden würden. Punkte, die diese Bedingung nicht erfüllen, liegen entweder innerhalb des zu rekonstruierenden Objekts oder sind Teil des zusätzlichen sogenannten *void*-Volumens, das bisher nicht eingesehen wurde und dessen Beseitigung eines der Ziele der Aufnahmeplanung ist. Daher soll die Funktion f so definiert werden, daß Punkte außerhalb des Objekts sich durch ihr Vorzeichen von den übrigen Punkten unterscheiden. Die Oberfläche des bisher eingesehenen Volumens wird damit durch die Nulldurchgänge der Funktion f definiert, während die vermessene Oberfläche durch $f(\mathbf{x}) = 0$ implizit definiert ist [73].

Um f zu definieren wird zunächst für jedes Tiefenbild r_i eine Distanzfunktion $g_i : \mathbb{R}^3 \rightarrow \mathbb{R}$ eingeführt, welche die vorzeichenbehaftete Distanz eines Punktes \mathbf{x} zur

¹Der geringe Speicherplatzbedarf wird unter der Annahme vollständig besetzter Tiefenbilder oft auch durch die Beschreibung "Verlustfreie Komprimierung um Faktor 12" auf den Punkt gebracht.

vermessenen Oberfläche entlang der Blickrichtung berechnet:

$$g_i(\mathbf{x}) := \begin{cases} r_i(u, v) - (0 \ 0 \ 1) \cdot h(R_i \mathbf{x} + t_i) & r_i(u, v) \text{ ist definiert} \\ -1 & \text{sonst} \end{cases} \quad (3.3)$$

Dabei ist (u, v) die Pixelposition, auf die \mathbf{x} durch h projiziert wird. Eine positive Distanz² zeigt also an, daß der betrachtete Punkt \mathbf{x} zwischen der vermessenen Oberfläche und dem Sensor, also außerhalb des Objekts, liegt. Aus den Werten $g_i(\mathbf{x})$ für mehrere Bilder kann nun eine konsistente Oberflächenbeschreibung für den Bildverbund durch

$$f(\mathbf{x}) := \max_i g_i(\mathbf{x}) \quad (3.4)$$

hergeleitet werden, die, wie Abbildung 3.2 zeigt, auch Eigenverdeckungen korrekt behandelt.

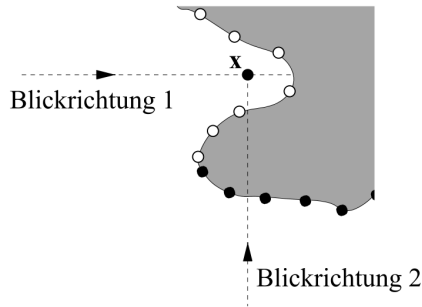


Abbildung 3.2: Hier ist $g_1(\mathbf{x}) > 0$, während $g_2(\mathbf{x}) < 0$ ist.

Im Falle von triangulationsbasierten Sensoren, die mehrere bekannte Beobachtungspunkte aufweisen, stellt sich die Frage, wie neben dem Blickpunkt des Tiefenbildes auch die übrigen Beobachtungspunkte zu jeder Aufnahme in die Definition der Funktionen f beziehungsweise g_i integriert werden können. Da die Berechnung von Schnittpunkten eines Sichtstrahls mit der vermessenen Oberfläche sehr aufwändig ist bietet es sich an, hierzu für jeden zusätzlichen Beobachtungspunkt ein Hilfsbild zu berechnen, welches durch Abbildung der Meßpunkte und Interpolation, gegebenenfalls mit reduzierter Auflösung, aus dem jeweiligen Tiefenbild gewonnen werden kann. Auf diese Weise kann die Schnittpunktberechnung auch für die zusätzlichen Blickpunkte direkt durch Vergleich des Pixelwerts an der Stelle des projizierten Punktes vereinfacht werden. Die zusätzlichen Hilfsbilder unterscheiden sich in ihrer Behandlung bei der Definition von f nicht von den primären Tiefenbildern.

²Ohne Beschränkung der Allgemeinheit wird hier davon ausgegangen, daß die gespeicherten Pixelwerte mit der gemessenen Distanz ansteigen.

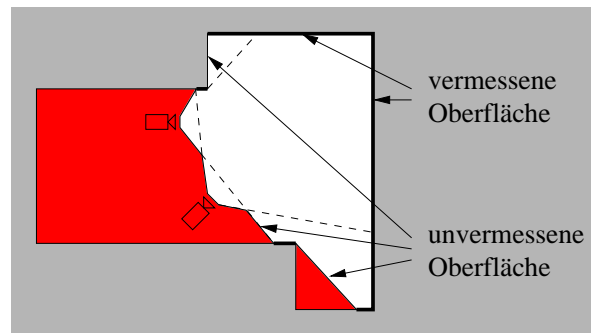


Abbildung 3.3: Der eingesehene Halbraum nach zwei Aufnahmen mit den vermessenen und interpolierten Oberflächenbereichen für die interne Rekonstruktion.

Die Definition der Funktion f ermöglicht prinzipiell die Klassifizierung beliebiger Punkte im Raum. In den meisten Fällen soll jedoch lediglich ein begrenztes Meßvolumen $V \subset \mathbb{R}^3$ in Betracht gezogen werden. So kann beispielsweise für Objekte meist ein umschließender Quader oder zumindest der Halbraum oberhalb der Standfläche angegeben werden. Auch im Falle von Innenräumen ist es oft sinnvoll, Grenzen des zu rekonstruierenden Bereichs anzugeben, wenn dieser etwa nicht physisch abgeschlossen ist (zum Beispiel durch eine offene Tür). Hier muß topologisch zwischen interner und externer Rekonstruktion unterschieden werden, da die Bereiche $x \notin V$ bei der internen Rekonstruktion innerhalb des zu rekonstruierenden Objekts liegen, bei der externen Rekonstruktion aber außerhalb. Mit entsprechenden Ergänzungen ergibt sich die Definition der Funktion g_i durch

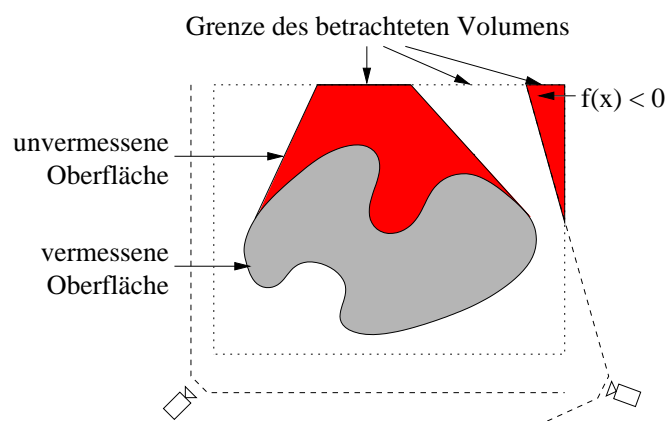


Abbildung 3.4: Situation wie in Abb. 3.3 für die externe Rekonstruktion, die zusätzlich die Grenze des betrachteten Volumens einbezieht.

$$g_i(\mathbf{x}) := \begin{cases} r_i(u, v) - (0 \ 0 \ 1) \cdot h(R_i \mathbf{x} + t_i) & \text{wenn } r_i(u, v) \text{ definiert und } \mathbf{x} \in V \\ +1 & \text{wenn } \mathbf{x} \notin V \text{ bei} \\ -1 & \text{externer Rekonstruktion} \\ & \text{sonst} \end{cases} \quad (3.5)$$

Die Nulldurchgänge von f definieren nun eine geschlossene und orientierbare Oberfläche S , die die bisher nicht eingesehenen Bereiche des Meßvolumens begrenzt. S ist eine 2-Mannigfaltigkeit und kann daher mit den üblichen Methoden der Computergraphik dargestellt werden. Das durch S begrenzte Volumen besteht zum einen aus dem Objektvolumen, zum anderen aus zusätzlichem *void*-Volumen, von dem zunächst nicht bekannt ist, ob und in welchem Umfang es weitere Teile des Objekts enthält. Die an das *void*-Volumen angrenzenden Teile der Oberfläche S , die durch eine Diskontinuität am Nulldurchgang von f gekennzeichnet sind, werden im Folgenden als unvermessene Teile von S oder einfach als *void*-Flächen bezeichnet.

Die Minimierung des *void*-Volumens ist eines der wesentlichen Ziele der Aufnahmeplanung. Da aber die Grenze zwischen *void*-Volumen und Objekt-Volumen nicht bekannt ist, kann dieses Ziel nicht direkt verfolgt werden. Die Strategie der Aufnahmeplanung besteht daher darin, zusätzliche Aufnahmen an den unvermessenen Bereichen von S auszurichten. Dabei kann insbesondere die Sichtbarkeitsbestimmung mit Hilfe von S durchgeführt werden, die zwar falsch-negative (berechnete Verdeckung ausschließlich durch *void*-Volumen verursacht), aber prinzipbedingt niemals falsch-positive Ergebnisse liefern kann. In den folgenden Abschnitten wird gezeigt, wie eine computergraphisch nutzbare Repräsentation von S effizient inkrementell konstruiert und mit weiteren Attributen, insbesondere solchen zur Repräsentation der Datenqualität, ergänzt werden kann.

3.4 Inkrementelle Konstruktion der Objektoberfläche

Die im vorigen Abschnitt beschriebene implizite Definition der Oberfläche S durch die Nulldurchgänge der Funktion f kann prinzipiell direkt zur Konstruktion einer computergraphisch darstellbaren Repräsentationsform genutzt werden. In der Literatur stehen zahlreiche Vorschläge zur Triangulierung implizit definierter Oberflächen zur Verfügung. Diese Verfahren berücksichtigen jedoch nicht die inkrementelle Natur der automatisierten 3D-Rekonstruktion, so daß die Repräsentation von S in jedem Schritt neu berechnet werden müßte. Tatsächlich kann durch Unterscheidung von vermessenen und unvermessenen Bereichen von S der Anteil neu zu berechnender Oberflächenteile deutlich reduziert werden.

Formal kann die Situation für den i -ten Aufnahmeschritt folgendermaßen beschrieben werden: Der Aufnahme i wird das vom Sensor eingesehene Volumen $v_i \subset \mathbb{R}^3$, die

vermessene Oberfläche $m_i \subset S$ sowie die unvermessene Oberfläche u_i des eingesehenen Volumens zugeordnet. Die Oberflächen $m_i \cup u_i$ bilden dabei die vollständige Begrenzung des eingesehenen Volumens v_i . Zur Integration mehrerer Aufnahmen in eine gemeinsame Oberflächenbeschreibung wird das gesamte in Schritt $1 \dots i$ eingesehene Volumen mit V_i , die in den Schritten $1 \dots i$ vermessene Oberfläche durch M_i , und die unvermessene Oberfläche des in den Aufnahmen $1 \dots i$ eingesehenen Volumens durch U_i bezeichnet. Offensichtlich sind initial $M_1 := m_1$, $U_1 := u_1$ und $V_1 := v_1$. Die Oberfläche S_i kann nun rekursiv durch

$$\begin{aligned} U_{i+1} &:= U_i \setminus v_{i+1} \cup u_{i+1} \setminus V_i \\ M_{i+1} &:= M_i \setminus v_{i+1} \cup m_{i+1} \setminus V_i \\ V_{i+1} &:= V_i \cup v_{i+1} \end{aligned} \quad (3.6)$$

definiert werden³. Um die Nutzung und weitere Vereinfachung dieses Ansatzes theoretisch abzusichern, wird im folgenden gezeigt, daß die beiden Beschreibungsformen tatsächlich die gleiche Oberfläche definieren.

SATZ 3.1 *Die inkrementelle Definition nach Gleichung 3.6 der Oberfläche*

$$S_i := U_i \cup M_i \quad (3.7)$$

ist äquivalent zur Definition der Oberfläche

$$S_i := \{\mathbf{x} \in \mathbb{R}^3 \mid \exists \delta \in \mathbb{R}^3 : f_i(\mathbf{x} + \delta) \cdot f_i(\mathbf{x} - \delta) < 0, |\delta| < \varepsilon\} \quad (3.8)$$

durch die Nulldurchgänge der Funktion f nach den Gleichungen 3.5 und 3.4 mit $\varepsilon \rightarrow 0$, wobei f_i die Funktion f bei Berücksichtigung der Aufnahmen $1 \dots i$ und der entsprechenden Funktionen $g_1 \dots g_i$ in Gleichung 3.4 bezeichnet.

BEWEIS 3.1 Für die initiale Konfiguration mit einer einzigen Aufnahme ist die Behauptung bereits durch die Definition von m_1 und u_1 erfüllt:

$$m_1 = \{\mathbf{x} \in \mathbb{R}^3 \mid f_1(\mathbf{x}) = 0\}, \quad (3.9)$$

wobei δ ein beliebig kurzer Vektor in der jeweiligen Blickrichtung des Sensors ist. Entsprechend ist

$$u_1 = \{\mathbf{x} \in \mathbb{R}^3 \mid \exists \delta \in \mathbb{R}^3 : f_1(\mathbf{x} + \delta) \cdot f_1(\mathbf{x} - \delta) < 0, f_1(\mathbf{x}) \neq 0\}, \quad (3.10)$$

³Die entsprechende Definition in den Konferenzbeiträgen [51, 53] enthält einen Notationsfehler, da die Zeichen für Differenz (\setminus) und Schnitt (\cap) versehentlich vertauscht wurden.

wobei δ ein beliebig kurzer Vektor in Richtung der Oberflächennormale von S ist. Ebenfalls nach Definition gilt

$$v_1 = \{\mathbf{x} \in \mathbb{R}^3 \mid f_1(\mathbf{x}) > 0\}. \quad (3.11)$$

Für den Übergang von S_i nach S_{i+1} ist nun zu zeigen, daß S_{i+1} einerseits alle zusätzlich durch die Berücksichtigung von g_{i+1} zu S_i hinzugekommenen Nulldurchgänge enthält, andererseits aber alle durch Berücksichtigung von g_{i+1} entfallenden Nulldurchgänge aus S_i nicht in S_{i+1} enthalten sind. Für zusätzliche Nulldurchgänge \mathbf{x}_0 muß

$$g_{i+1}(\mathbf{x}_0) = 0 \quad (3.12)$$

(in Aufnahme $i + 1$ vermessene Oberfläche) oder

$$g_{i+1}(\mathbf{x}_0 + \delta) \cdot g_{i+1}(\mathbf{x}_0 - \delta) < 0 \vee g_{i+1}(\mathbf{x}_0) \neq 0 \quad (3.13)$$

(unvermessene Grenzfläche des in Aufnahme $i + 1$ eingesehenen Volumens) erfüllt sein; wegen der Maximum-Bildung in Gleichung 3.4 muß zusätzlich

$$f_i(\mathbf{x}_0) \leq 0 \quad (3.14)$$

erfüllt sein. Für die entfallenden Nulldurchgänge \mathbf{x}_0 gilt ebenfalls wegen der Maximum-Bildung in Gleichung 3.4

$$g_{i+1}(\mathbf{x}_0) > 0 \quad (3.15)$$

Nach der Definition von u_i , m_i und v_i gelten folgende Äquivalenzrelationen, die offensichtlich den in den Gleichungen 3.12, 3.13 und 3.15 beschriebenen Fällen entsprechen:

$$u_{i+1} = \{\mathbf{x} \in \mathbb{R}^3 \mid g_{i+1}(\mathbf{x} + \delta) \cdot g_{i+1}(\mathbf{x} - \delta) < 0, g_{i+1}(\mathbf{x}) \neq 0\} \quad (3.16)$$

$$m_{i+1} = \{\mathbf{x} \in \mathbb{R}^3 \mid g_{i+1}(\mathbf{x}) = 0\} \quad (3.17)$$

$$v_{i+1} = \{\mathbf{x} \in \mathbb{R}^3 \mid g_{i+1}(\mathbf{x}) > 0\} \quad (3.18)$$

Zusammen mit den Gleichungen zur intialen Konfiguration 3.9, 3.10 und 3.11 ergibt sich durch Einsetzen in Gleichung 3.6 und vollständige Induktion die Behauptung.

q.e.d.

Zur Vereinfachung der Konstruktion der Oberfläche S ist es sinnvoll, die einzelnen Elemente in Gleichung 3.6 auf ihre geometrische Relevanz hin zu untersuchen. Während für die unvermessene Grenzfläche des eingesehenen Volumens U_i durch die Differenzbildung mit dem eingesehenen Volumen erhebliche Geometrieänderungen zu

erwarten sind, ist dies bei der Kombination der vermessenen Oberfläche nicht der Fall. Hier ist die Differenzbildung mit dem eingesehenen Volumen im wesentlichen erforderlich, um eine geschlossene 2-Mannigfaltigkeit zu erhalten; die durch diese Operation verursachten geometrischen Unterschiede liegen in der Größenordnung der Meßgenauigkeit des Sensors. An dieser Stelle bestehen daher bei der Implementierung Freiheitsgrade, um zugunsten des Rechenaufwands eine geringfügige geometrische Ungenauigkeit zu tolerieren. Zwei derartige Ansätze, die im Rahmen der vorliegenden Arbeit verwendet wurden, werden im Folgenden vorgestellt, während die Konstruktion der Grenzfläche des eingesehenen Volumens U_i Gegenstand von Abschnitt 3.4.3 ist.

3.4.1 Vermaschung überlappender Dreiecksnetze

Das Verfahren der Vermaschung von Dreiecksnetzen (engl. *zippered polygon meshes* [110]) kombiniert die teilvermessenen Oberflächen aus mehreren Tiefenbildaufnahmen, indem jeder in Aufnahme $i + 1$ vermessene Oberflächenabschnitt m_{i+1} zuerst einzeln trianguliert und anschließend mit dem in den vorangegangenen Schritten $1 \dots i$ erhaltenen Dreiecksnetz T_i verbunden werden. Hierzu können in einer für die Aufnahmeplanung optimierten Variante [99] des Verfahrens alle bereits in T_i repräsentierten Bereiche von m_{i+1} bei der Konstruktion des Dreiecksnetzes t_{i+1} des vermessenen Oberflächenabschnitts m_{i+1} weggelassen werden, so daß im Ergebnis immer die zuerst erfolgte Aufnahme Eingang in das Dreiecksnetz T_i findet. Die korrespondierenden Ränder von T_i und t_{i+1} werden durch zusätzliche Dreieckskanten miteinander verbunden (*“zippering”*); die Oberflächennormale kann für jeden Dreieckseckpunkt aus den angrenzenden Dreiecken gemittelt werden.

Gegenüber der direkten Umsetzung der Definition in Gleichung 3.6, die ein *clipping* der Dreiecksnetze t_{i+1} und T_i gegen das eingesehene Volumen V_i beziehungsweise v_{i+1} erfordern würde, ist hier lediglich die Überlappung von t_{i+1} mit M_i zu prüfen. Diese Operation kann für alle Eckpunkte von t_{i+1} effizient durch Projektion in die bisher aufgenommenen Tiefenbilder ausgeführt werden. Insbesondere ist es nicht erforderlich, korrespondierende Punkte auf dem Dreiecksnetz T_i zu suchen, so daß aufwändige Datenstrukturen zur Unterstützung dieser Suche wie *binary space partitioning trees* vermieden werden können. Auch der Aufwand für das initiale Auffinden korrespondierender Punkte auf den Rändern von T_i und t_{i+1} läßt sich durch Nutzung der Projektionsfunktion unabhängig von der Anzahl der beteiligten Dreieckseckpunkte halten. Daher ist die Komplexität des Verfahrens für die Integration eines zusätzlichen Bildes linear in der Anzahl der bisher akquirierten Aufnahmen.

Der offensichtlichste Nachteil dieser Vorgehensweise ist der Erhalt qualitativ suboptimal repräsentierter Teile der Oberfläche, der durch die ausschließliche Berücksichtigung bisher unvermessener Teile von m_{i+1} in T_{i+1} zustande kommt. Wie in Abschnitt 3.6 deutlich wird, reicht es aber aus, eine topologisch korrekte (das heißt: alle

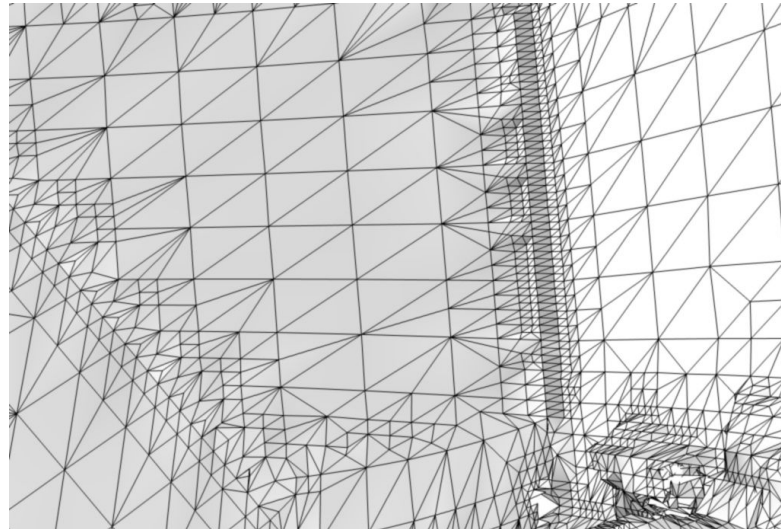


Abbildung 3.5: Vereinigung der Dreiecksnetze aus drei beteiligten Einzelaufnahmen durch mesh zippering. Durch die uniforme Größe und die bildachsparallele Anordnung ist der Übergang zwischen den Aufnahmen auf der planaren Fläche links im Bild deutlich zu erkennen.

Löcher und Henkel repräsentierende) Initialtriangulierung mit einer gewissen geometrischen Mindestgenauigkeit zu bestimmen. Geometrische Ungenauigkeiten können nachträglich durch Projektion der Dreieckseckpunkte in Richtung der Oberflächennormalen auf die implizit durch alle Tiefenbildaufnahmen definierte Oberfläche $f(\mathbf{x}) = 0$ beseitigt werden, während der Interpolationsfehler des Dreiecksnetzes durch das adaptive Einfügen zusätzlicher Dreieckseckpunkte auf einen vorgegebenen Schwellwert reduziert werden kann [73]. Für die Aufnahmeplanung ist die so erreichbare Genauigkeit allerdings im Regelfall gar nicht erforderlich, so daß diese Operation nur einmalig am Ende des Rekonstruktionsprozesses durchgeführt werden muß.

3.4.2 Punktrendering

Ein Nachteil der Repräsentation von S durch ein Dreiecksnetz besteht in der konstanten Auflösung, die nicht an die Bildauflösung bei der gerenderten Darstellung angepaßt werden kann. Wie bereits am Anfang dieses Kapitels angemerkt wurde, sollen aber wesentliche Teile der Aufnahmeplanung, insbesondere die Sichtbarkeitsbestimmung, aus Gründen des Rechenaufwands nur mit der Genauigkeit eines relativ niedrig aufgelösten Bildrasters berechnet werden. Hier besteht die Gefahr, daß ein hochauflöstes Dreiecksnetz bei der Projektion in das Bildraster pro Dreieck eine Fläche belegt, die wesentlich kleiner als ein Pixel ist. Auf diese Weise bleibt ein erheblicher

Teil der im Dreiecksnetz enthaltenen Detailinformation ungenutzt. Im Folgenden wird eine Alternative zum klassischen Ansatz der Dreiecksnetzvereinfachung vorgestellt, der die Oberfläche durch punktförmige Elemente darstellt. Die Schwierigkeit besteht dabei darin, für eine gegebene Punktmenge, einen gegebenen Blickpunkt und für eine gegebene Projektionsvorschrift die einem Punkt zugeordnete Bildfläche so zu wählen, daß keine Lücken zwischen den flächenhaft dargestellten Punktelementen auftreten.

Da die darzustellende Oberfläche nativ durch die Tiefenbild-Meßwerte punktweise definiert ist, erfordert dieser Ansatz nur geringen Aufwand zur Aufbereitung der Daten. Zusätzlich kann die geordnete Repräsentation der Meßpunkte im Bildraster genutzt werden, um weitere Eigenschaften der Oberfläche, insbesondere die Oberflächennormale und die lokale Abtastdichte, effizient durch lokale Operatoren aus dem Tiefenbild zu berechnen, so daß die entsprechenden aufwändigen Rechenschritte in der Literatur vorgeschlagener Verfahren [2] hier vermieden werden können. Die Berücksichtigung der Oberflächennormale ermöglicht die schattierte computergraphische Darstellung; dies ist eine wichtige Voraussetzung für das in Kapitel 4 vorgeschlagene Prädiktionsverfahren.

Im Vergleich zur Definition der Oberfläche M_i in Gleichung 3.6 erfolgt hier die Kombination mehrerer Aufnahmen durch die einfache Vorschrift

$$M_{i+1} := M_i \cup m_{i+1}, \quad (3.19)$$

wodurch die 2-Mannigfaltigkeits-Bedingung der entstehenden Oberfläche verletzt wird, da mehrere hintereinanderliegende und sich potentiell durchdringende Teiloberflächen in den überlappend vermessenen Bereichen entstehen. Welche Teiloberflächen m_i in der graphischen Darstellung von M_i tatsächlich als sichtbar klassifiziert und angezeigt werden, hängt in diesen Bereichen vom Blickpunkt, von der *z-buffer*-Quantisierung und von der Reihenfolge ab. Im Unterschied zur Darstellung durch ein Dreiecksnetz, die bei der in Abschnitt 3.4.1 beschriebenen Konstruktion abschnittsweise stetige Oberflächen ergibt, liefert die Darstellung durch Punktrendering unstetige Oberflächen, da der Abstand des jeweils dargestellten Punktelements zum Blickpunkt im Rahmen der Streuung der Tiefenbildmeßwerte pixelweise variieren kann. Daher verbietet sich bei der Darstellung durch Punktrendering die Auswertung der Oberflächennormale des computergraphischen *z-buffers*, die bei der Darstellung durch ein Dreiecksnetz dagegen möglich ist.

Zusätzlich zum (in diesem Fall fast trivialen) Aufbau der Oberflächenrepräsentation durch die in Gleichung 3.19 definierte Kombination der darzustellenden Punktelemente erfordert das Punktrendering die Wahl der Größe der einem Punktelement zugeordneten Bildfläche. Hierzu muß zunächst die minimale von einem Bildpunkt abgedeckte Fläche A_s auf der darzustellenden Oberfläche bekannt sein, die sich für einen minimalen Blickpunkt Abstand z_{min} , einen perspektivischen Öffnungswinkel α und eine durch

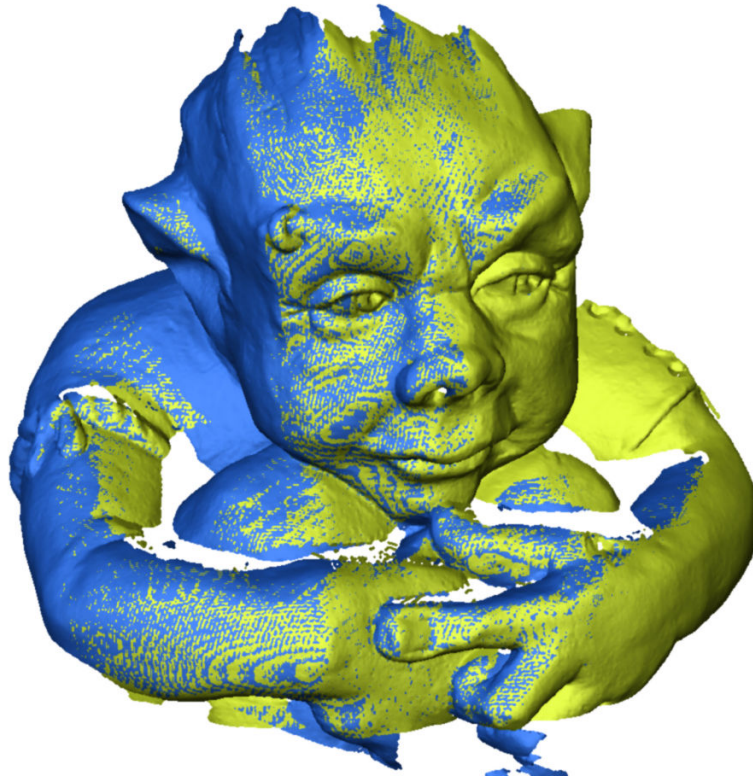


Abbildung 3.6: Vereinigung der Daten der ersten zwei Aufnahmen eines Beispielobjekts durch Punktrendering. Die für die Schattierung verwendeten Oberflächennormalen werden mit lokalen Operatoren aus den Tiefenbildern bestimmt.

α abgedeckte Bildauflösung von s Rasterpunkten durch

$$A_s = \left(z_{\min} \cdot \frac{\tan \alpha}{2 \cdot s} \right)^2 \quad (3.20)$$

abschätzen läßt. Für eine minimal auftretende Abtastdichte von d Meßpunkten pro Flächeneinheit kann die zur lückenlosen Darstellung erforderliche Größe A_p eines Punktelements nun durch

$$A_p = \frac{1}{d \cdot A_s} \quad (3.21)$$

angegeben werden. Daher muß die minimal auftretende Abtastdichte bei der Integration einer zusätzlichen Aufnahme unter Berücksichtigung mehrfach vermessener Bereiche neu bestimmt werden. Zur Vermeidung der ineffizienten Mehrfach-Darstellung der Bildpunkte im Fall $A_p < 1$ kann die Anzahl der tatsächlich zur Anzeige gebrachten Punktelemente reduziert werden, so daß sich der Wert d entsprechend verkleinert. Da die lokale Abtastdichte zu jedem Punkt der Oberfläche S ohnehin für die Auswertung eines Qualitäts-Kriteriums bei der Aufnahmeplanung benötigt wird, fällt an dieser

Stelle kein zusätzlicher Rechenaufwand an. Die Berechnung der Abtastdichte wird in Abschnitt 3.4.4 genauer dargestellt. Da dieser Schritt auch den Rechenaufwand des Punktrenderings in Abhängigkeit von der Bildanzahl begrenzt, wird zur Betrachtung der Komplexität des Verfahrens ebenfalls auf Abschnitt 3.4.4 verwiesen.

3.4.3 Grenzflächen des eingesehenen Volumens

Durch die geradlinige Ausbreitung der Sichtstrahlen ist der unvermessene Rand U_i des eingesehenen Volumens V_i zumindest in einer der Hauptkrümmungsrichtungen ungekrümmt und kann daher effizient durch ein Dreiecksnetz repräsentiert werden. Im Folgenden wird beschrieben, wie die Definition aus Gleichung 3.6 bei der inkrementellen Konstruktion einer Repräsentation von U_i direkt umgesetzt werden kann. Anders als im Beweis zu Satz 3.1 wird hier jedoch der Rand des betrachteten Volumens gesondert berücksichtigt.

Für jede Aufnahme r_i werden zunächst an den Sprungkanten die eigentlich nicht zur Interpolation geeigneten Datenpunkte durch Dreiecke verbunden. Zusätzlich wird aus den Rändern des definierten Bereichs von r_i durch eine *sweep*-Operation in Richtung der jeweiligen Sensor-Blickrichtung eine seitliche Begrenzungsfläche gewonnen, die im Falle einer innerhalb des betrachteten Volumens liegenden Sensorposition sensorseitig geschlossen wird, andernfalls aber am Rand des betrachteten Volumens offen bleibt. Das entstehende, im Allgemeinen aus mehreren Teilen bestehende Dreiecksnetz wird mit t_i bezeichnet.

Die Berücksichtigung des Randes des betrachteten Volumens erfordert für die interne und externe Rekonstruktion ein weiteres Dreiecksnetz t_v , das die zusätzliche Grenzfläche repräsentiert. Dieses kann für die externe Rekonstruktion analog zu den übrigen t_i behandelt werden, indem die zugehörige Oberfläche u_v durch

$$U_1 := u_v \setminus v_1 \cup u_1 \quad (3.22)$$

initial berücksichtigt wird. Bei der internen Rekonstruktion ist die Grenze des betrachteten Volumens initial kein Teil von S , sondern wird erst durch einen Schnitt mit dem eingesehenen Volumen Teil von S . Daher muß u_v in diesem Fall initial durch

$$U_1 := u_v \cap v_1 \cup u_1 \quad (3.23)$$

und in jedem weiteren Iterationsschritt durch

$$U_{i+1} := U_i \setminus v_{i+1} \cup u_{i+1} \setminus V_i \cup u_v \cap v_{i+1} \quad (3.24)$$

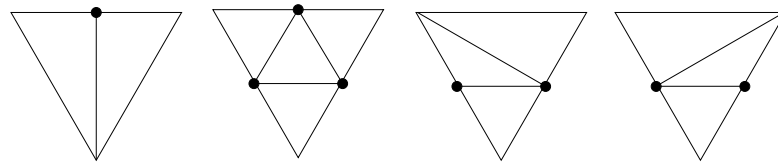


Abbildung 3.7: Unterteilungsschema für das clipping der Dreiecksnetze.

berücksichtigt werden. Die Differenz der Flächen u_i gegen das betrachtete Volumen wird sowohl bei der internen wie bei der externen Rekonstruktion bereits bei der Konstruktion der Repräsentation t_i von u_i berücksichtigt.

In allen Fällen können Differenz- und Schnitt-Operation auf dem Dreiecksnetz durch einen einfachen *clipping*-Algorithmus implementiert werden. Dabei werden für jedes Dreieck drei Fälle unterschieden:

1. Dreiecke, die vollständig im eingesehenen Volumen einer Aufnahme liegen, werden verworfen.
2. Dreiecke, die vollständig außerhalb des in den bisherigen Aufnahmen eingesehenen Volumens liegen, bleiben unverändert erhalten.
3. Alle übrigen Dreiecke werden mit einem einfachen, in Abbildung 3.7 dargestellten Schema unterteilt.

Dies wird wiederholt, bis alle zu unterteilenden Dreiecke eine gewisse Mindestgröße unterschreiten. Als Unterteilungskriterium wird jeweils für die beiden Dreieckseckpunkte einer Kante geprüft, ob die Zugehörigkeit zum eingesehenen Volumen verschieden ist; in diesem Fall kann auch ein optimierter Unterteilungspunkt durch *bisection*-Suche gewonnen werden. Ist dieses Kriterium für keine der Kanten erfüllt, so werden alle drei Kanten am jeweiligen Kantenschwerpunkt unterteilt. Nach Anwendung dieses Algorithmus' wird das Dreiecksnetz auf Basis einer Krümmungsanalyse vereinfacht.

Die im Allgemeinen nicht zusammenhängenden Teile des entstehenden Dreiecksnetzes können an den korrespondierenden Rändern durch Einfügen zusätzlicher Kanten verbunden werden. Sofern die Lücken zwischen den Teilen des Dreiecksnetzes klein sind, ist dies für die Aufnahmeplanung zunächst nicht notwendig. Wenn allerdings auch die vermessene Oberfläche durch vermaschte Dreiecksnetze repräsentiert ist, so kann durch das Verbinden der Ränder eine geschlossene („wasserdichte“) polygonale Repräsentation der Oberfläche S gewonnen werden, so daß bei der späteren Verarbeitung des vollständigen Datensatzes, wie sie in Abschnitt 3.6 skizziert wird, der Aufbau einer entsprechenden Initialtriangulierung nicht mehr notwendig ist.

Die Komplexität des Verfahrens ist für die Integration eines zusätzlichen Bildes offensichtlich linear in der Anzahl der bisher akquirierten Aufnahmen, da die Dreiecksnetze zu allen bisherigen Aufnahmen gegen das neu eingesehene Volumen geschnitten

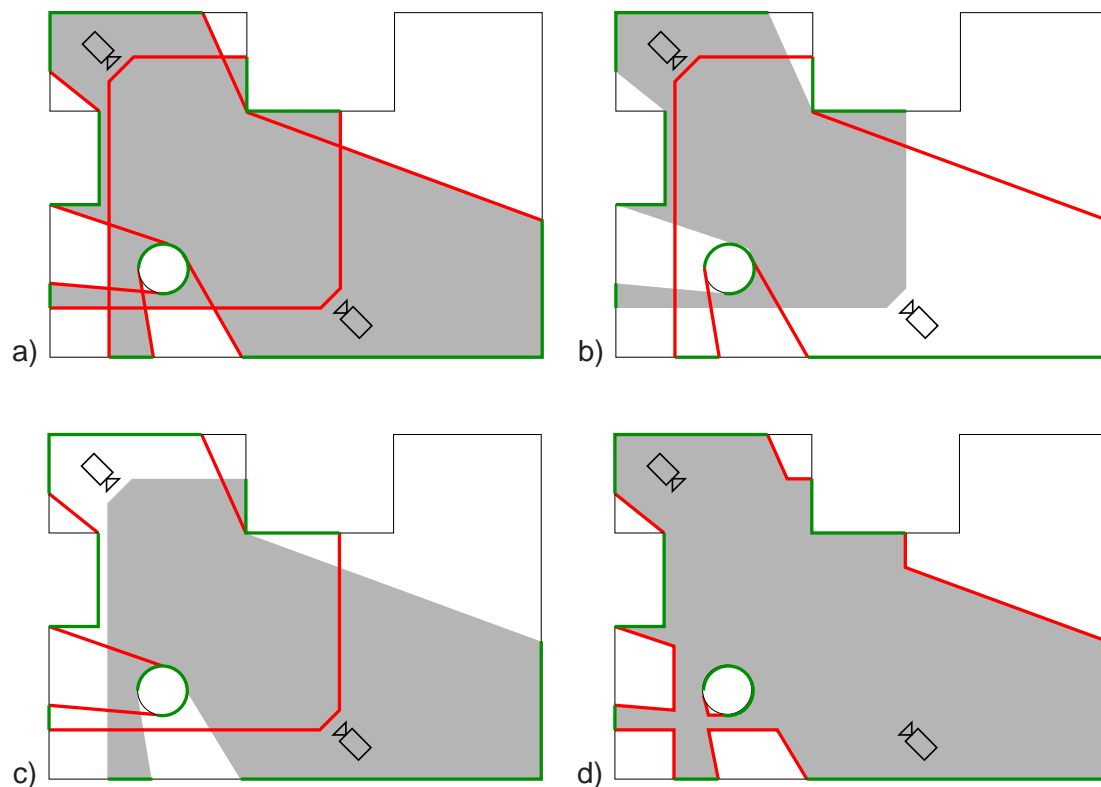


Abbildung 3.8: Die unvermessenen Grenzflächen (rot) der beiden Aufnahmen in a) werden durch Clipping gegen das eingesehene Volumen (grau) vereinigt. In b) ist der Schnitt der neuen Grenzflächen gegen das eingesehene Volumen der vorherigen Aufnahme dargestellt, in c) der Schnitt der vorhandenen Grenzflächen gegen das neu eingesehene Volumen. Die vermessenen Oberflächen (grün) werden vereinigt. Das Ergebnis ist in d) zu sehen.

werden müssen, während das neu hinzuzufügende Dreiecksnetz gegen das in allen bisherigen Aufnahmen eingesehene Volumen geschnitten werden muß.

3.4.4 Repräsentation qualitativer Maßzahlen

Neben der Geometrie der Objektrepräsentation und der Möglichkeit, unvermessene Grenzflächen des eingesehenen Volumens zu identifizieren, muß auch die Qualität der verfügbaren Daten zu den bereits vermessenen Bereichen der Oberfläche in der Objektrepräsentation enthalten sein. Die naheliegendste qualitative Maßzahl, die einem Oberflächenpunkt zugeordnet werden kann, ist die lokale Abtastdichte; sie ist gleichzeitig ein häufig vorgegebenes Qualitätsmaß für die 3D-Digitalisierung. Im Falle bereits vermessener Oberflächen kann außerdem eine adaptive, das heißt, lokal unter-

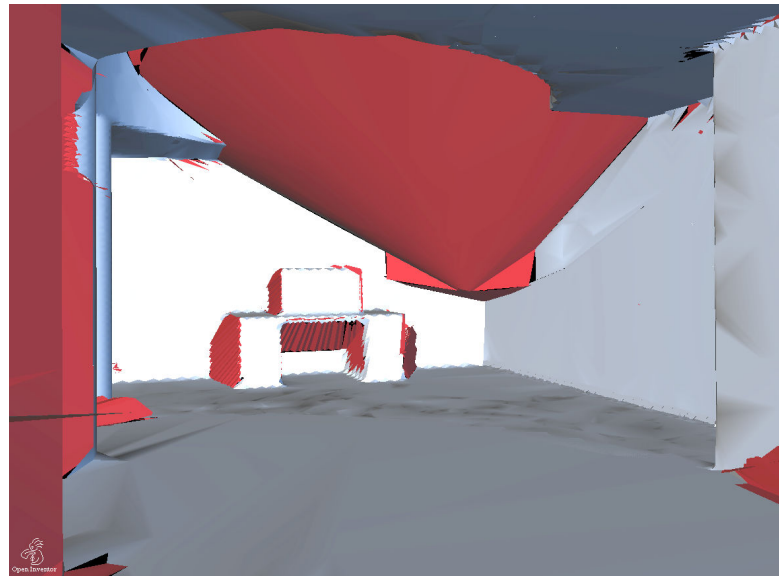


Abbildung 3.9: Grenzflächen des eingesehenen Volumens (rot) basierend auf simulierten Aufnahmen einer künstlich definierten Testszene. Die Bildgeometrie des simulierten 3D-Sensors ist im wesentlichen sphärisch.

schiedliche Soll-Abtastdichte vom Benutzer vorgegeben werden, um beispielsweise für wichtige Details eine erhöhte Auflösung zu spezifizieren. Weitere mögliche Maßzahlen sind (ohne Anspruch auf Vollständigkeit) der flachste Winkel, aus dem die vermessene Oberfläche von einem der an der Messung beteiligten Blickpunkte aus eingesehen wurde, der längste und der kürzeste Meßabstand der bisherigen Messungen, sowie die Varianz der einzelnen Messungen in mehrfach vermessenen Bereichen.

Zur Begrenzung des entstehenden Rechenaufwands erscheint es sinnvoll, sich auf wenige Maßzahlen zu beschränken, damit diese computergraphisch durch die übliche dreikanalige Farbdefinition (RGB) repräsentiert werden können. Hierzu wird im Folgenden zunächst ein Qualitätsmaß eines Meßpunktes entwickelt, das im Falle von triangulationsbasierten Stereo-Sensoren alle beteiligten Blickpunkte des Sensors berücksichtigt. Bei der Aufnahmeplanung soll diese Maßzahl verwendet werden, um den Qualitätsgewinn einer zusätzlichen Aufnahme gegenüber zuvor erfolgten Messungen zu quantifizieren.

3.4.4.1 Laterale Abtastdichte

Die Abtastdichte eines Punktes $p_w \in \mathbb{R}^3$ in Weltkoordinaten in einer Aufnahme i läßt sich auf einfache Weise mit Hilfe der in Gleichung 3.2 eingeführten Projektionsfunktion h berechnen, indem an der Stelle $p_r = h(R_i \cdot p_w + t_i)$ zunächst zwei Tangentialvektoren

\mathbf{t}_u und \mathbf{t}_v im Bildraster berechnet werden durch

$$\mathbf{t}_u = \begin{pmatrix} 1 \\ 0 \\ \left. \frac{\partial r_i(u,v)}{\partial u} \right|_{\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} p_r} \end{pmatrix} \text{ und } \mathbf{t}_v = \begin{pmatrix} 0 \\ 1 \\ \left. \frac{\partial r_i(u,v)}{\partial v} \right|_{\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} p_r} \end{pmatrix} \quad (3.25)$$

Diese besitzen projiziert auf die Bildebene offensichtlich gerade den Abstand zweier Abtastpunkte als Länge. Da die Skalierung von Weltkoordinaten und Kamerakoordinaten identisch ist, genügt es, diese Vektoren in Kamerakoordinaten zu transformieren durch Multiplikation mit der Jacobimatrix der Funktion h^{-1} an der Stelle p_r , um mit Hilfe des Vektorprodukts die Fläche des von diesen Vektoren aufgespannten Parallelogramms in Weltkoordinaten zu berechnen. Die Abtastdichte d ergibt sich damit durch

$$d = \left\| \left(\left. \frac{\partial h^{-1}(x)}{\partial x} \right|_{x=p_r} \mathbf{t}_u \right) \times \left(\left. \frac{\partial h^{-1}(x)}{\partial x} \right|_{x=p_r} \mathbf{t}_v \right) \right\|^{-1} \quad (3.26)$$

3.4.4.2 Entfernungsmeßgenauigkeit

Im Falle eines unifokalen Sensors läßt sich die Abtastdichte direkt als Qualitätsmaß einsetzen, während im Falle eines triangulationsbasierten Stereo-Sensors auch die Ortsauflösung der Oberfläche bezüglich der übrigen Blickpunkte berücksichtigt werden muß. Im Gegensatz zum primären Tiefenbild trägt allerdings die Auflösung bezüglich der sekundären Blickpunkte nicht zur lateralen Ortsauflösung, sondern lediglich zur Genauigkeit des gespeicherten Tiefenwertes bei. Daher genügt es, die Auflösung bezüglich eines sekundären Blickpunktes eindimensional auf der Schnittkurve der Epipolar-Ebene mit der Objektoberfläche zu berechnen.

Um diese Größe auch ohne die explizite Speicherung des Sekundärbildes berechnen zu können, muß zunächst die Oberflächennormale \mathbf{n} durch

$$\mathbf{n} = \left(\left. \frac{\partial h^{-1}(x)}{\partial x} \right|_{x=p_r} \mathbf{t}_u \right) \times \left(\left. \frac{\partial h^{-1}(x)}{\partial x} \right|_{x=p_r} \mathbf{t}_v \right) \quad (3.27)$$

sowie die Sensorblickrichtung durch

$$\mathbf{c} = \left. \frac{\partial h^{-1}(u, v, w)}{\partial w} \right|_{\begin{pmatrix} u \\ v \\ w \end{pmatrix} = p_r} \quad (3.28)$$

aus dem primären Bild bestimmt werden. Mit der relativen Orientierung (R_{sek}, t_{sek}) der sekundären Kamera, dem transformierten betrachteten Punkt $p_{sek} = R_{sek}(R_i \cdot p_w +$

$t_i) + t_{sek}$ und der zugehörigen Projektionsfunktion h_{sek} können die entsprechenden Größen im Bildraster des sekundären Bildes durch

$$p'_r = h_{sek}(p_{sek}) \quad (3.29)$$

$$\mathbf{c}'_r = \left. \frac{\partial h_{sek}(x)}{\partial x} \right|_{x=p_{sek}} \frac{R_{sek} \cdot \mathbf{c}}{\|\mathbf{c}\|} \quad (3.30)$$

sowie die sekundäre Blickrichtung und die transformierte Oberflächennormale mit

$$\mathbf{c}_{sek} = \left. \frac{\partial h_{sek}^{-1}(u, v, w)}{\partial w} \right|_{\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \mathbf{p}'_r} \quad (3.31)$$

$$\mathbf{n}' = R_{sek} \cdot \mathbf{n} \quad (3.32)$$

bestimmt werden. Der gesuchte Auflösungswert d' kann nun durch

$$d' = \left\| \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \mathbf{c}'_r \right\| \cdot \frac{\mathbf{c}_{sek} \cdot \mathbf{n}'}{\|\mathbf{c}_{sek}\| \cdot \|\mathbf{n}'\|} \quad (3.33)$$

berechnet werden, indem zuerst die sekundäre laterale Auflösung entlang des Sichtstrahls der primären Aufnahme berücksichtigt wird, die dann mit $\mathbf{c}_{sek} \cdot \mathbf{n}'$ durch einen Kosinus-Zusammenhang zum Winkel zwischen dem Sichtstrahl der sekundären Kamera und der vermessenen Oberfläche in Beziehung gesetzt wird. Dabei geht der Triangulationswinkel des Sensors bereits durch die Projektion des normierten Sichtstrahls der primären Kamera in Gleichung 3.30 in die Rechnung ein.

Durch die allgemein gehaltene Formulierung auf Basis der Projektionsfunktionen h und h_{sek} wird auch ein wichtiger Sonderfall triangulationsbasierter Sensoren abgedeckt, der bei der eindimensionalen Musterprojektion (Streifenlichtprojektion durch spezielle Projektoren oder Laserlichtschnitt) auftritt. Dieses Verfahren führt bei der Kalibrierung des Projektors zu einem projektiven System, dessen Rasterauflösung in der einen Bildachse einer Kamera vergleichbar ist, in der anderen aber nahezu Null ist. Typischerweise wird ein derartiger Projektor so zur Kamera angeordnet, daß die Sichtstrahlen der Kamera, projiziert in das Bildraster des Projektors, möglichst parallel zur hochauflösenden Bildachse verlaufen, und die niedrige Auflösung der anderen Bildachse keinen Einfluß auf die Meßgenauigkeit hat.

3.4.4.3 Kombiniertes Qualitätsmaß

Der erhaltene flächenbezogene Auflösungswert d und die zusätzlichen, auf den Tiefenwert bezogenen Maßzahlen $d', d'', \dots, d^{(n)}$ sollen nun zu einer einzigen Maßzahl

kombiniert werden. Dazu werden zunächst die tiefenwertbezogenen Werte durch ihr geometrisches Mittel

$$d_r = \sqrt[n]{\prod_{i=1}^n d^{(i)}} \quad (3.34)$$

zusammengefaßt. Abhängig von der tatsächlichen Sensorkonfiguration sollte natürlich ein Kombinationsverfahren verwendet werden, das den tatsächlichen Meßvorgang exakter modelliert. Die hier vorgeschlagene Methode wurde für den Fall konzipiert, daß die dafür nötigen Detailkenntnisse nicht vorliegen. Ziel ist es daher, die Genauigkeit bereits bei Verwendung eines einzigen suboptimalen Blickpunktes deutlich niedriger zu bewerten, da die zusätzlichen Blickpunkte meist nicht kumulativ in Form redundanter Mehrfachmessung, sondern zur Triangulation mit dem primären Blickpunkt verwendet werden. Aus d und d_r wird nun durch Multiplikation eine volumenbezogene Auflösungsmaßzahl d_s berechnet, die die Dichte potentieller Meßpunkte pro Volumeneinheit am jeweiligen Punkt der Oberfläche angibt.

3.4.4.4 Kombination in mehrfach vermessenen Bereichen

Der Qualitätsgewinn, der sich durch die erneute Vermessung eines bereits erfaßten Oberflächenbereichs ergibt, hängt naturgemäß stark von den Details der weiteren Nutzung der Daten sowie von den Abtast-Eigenschaften des Sensors ab. Für letzteren Aspekt ist der Füllfaktor der Bildpunkte eine wichtige Maßzahl. Dieser gibt an, welcher Anteil des einem Bildpunkt zugeordneten Raumwinkels tatsächlich zur Bestimmung des Meßwertes genutzt wird; typischerweise findet dabei eine Integration über diesen Raumwinkel statt. Insbesondere bei der laserbasierten Lichtlaufzeitmessung kann der Füllfaktor erheblich in beide Richtungen von 100% abweichen und ist oftmals entfernungsabhängig, während bei kamerabasierten Systemen meist konstante Füllfaktoren knapp unter 100% anzutreffen sind.

Nur bei kleinen Füllfaktoren $\ll 100\%$ ergibt sich durch eine zusätzliche Aufnahme eine potentielle Steigerung der flächenbezogenen Auflösung, da auf diese Weise, anschaulich betrachtet, „Lücken“ zwischen den gemessenen Flächenstücken geschlossen werden. In allen anderen Fällen wird lediglich die Größe der Stichprobe erhöht, aus der mit statistischen Methoden die tatsächliche Lage der Oberfläche geschätzt werden kann. Beide Fälle erfordern spezialisierte Verfahren, um den möglichen Qualitätsgewinn tatsächlich zu realisieren.

3.4.4.5 Inkrementelle Berechnung

Die Auflösungsmaßzahl d_s wird mit den Geometrieelementen der Objektrepräsentation, im Falle vermaschter Dreiecksnetze also mit den Dreieckseckpunkten, im Falle

des Punktrenderings mit den Punktelementen, gespeichert. Initial kann sie für jedes Geometrieelement lokal im Tiefenbild berechnet werden. Für neu aufgenommene Tiefenbilder muß nun jedes Geometrieelement in alle bisher aufgenommenen Bilder projiziert werden, während die Geometrieelemente zu den übrigen Aufnahmen in das neu aufgenommene Tiefenbild projiziert werden. Dabei wird jeweils eine unterschiedliche Anzahl n von Korrespondenzen gefunden, zu denen die zugehörigen Werte $d_r^{(j)}$ und $\max_j d(j)$ für $1 \leq j \leq n$ bestimmt werden.

Abbildung 3.10 zeigt die Datensituation für mehrere Einzelaufnahmen, in denen die Qualität der Daten auf der Objektoberfläche farbkodiert dargestellt wird.



Abbildung 3.10: Kombinierte Auflösungsmaßzahl für eine Mehrbildsituation aus 5 Einzelaufnahmen, die hier farbkodiert auf der Objektoberfläche dargestellt ist. Grüne Bereiche besitzen eine hohe Abtastdichte, rote eine niedrige.

3.5 Initiale Repräsentation

Durch die Definition der Problemstellung, die von initial unbekannten Objektgeometrien ausgeht, liegen zu Beginn der automatisierten Rekonstruktion keine Tiefenbilddaten

der zu rekonstruierenden Oberfläche vor. Dennoch gibt es fast immer andere Informationen bezüglich der relativen Lage des Sensors zum Objekt, die initial berücksichtigt und in die Objektrepräsentation integriert werden können.

3.5.1 Interne Rekonstruktion

Im Falle der internen Rekonstruktion kann zunächst davon ausgegangen werden, daß sich der Sensor im zu rekonstruierenden Raum befindet. Zusätzlich ist seine absolute Höhe bekannt; die daraus abgeleitete Lage des ebenen Bodens stellt in vielen Fällen auch ein wichtiges Kriterium zur Erreichbarkeit potentieller Aufnahmepositionen durch den mobilen Roboter, der den Sensor trägt, dar. Ebenso ist die vom mobilen Roboter eingenommene Standfläche bekannt; das entsprechende Volumen kann kein Teil des zu rekonstruierenden Objekts sein. Daher kann initial ein zusätzliches „quasi-eingesehenes“ Volumen angenommen werden.

Die unvermessene Oberfläche dieses zusätzlichen eingesehenen Volumens ist allerdings in Bezug auf die Planung der ersten Aufnahme weitgehend aussageelos, da sie nur durch die Höhe des Sensors über dem Boden mit der Geometrie der zu vermessenden Oberfläche korreliert. Daher wird das initial „quasi-eingesehene“ Volumen v_0 und die zugehörige Oberfläche u_0 erst bei der Integration der ersten Aufnahme berücksichtigt, indem die Definition aus den Gleichungen 3.6, 3.23 und 3.24 zu

$$U_1 := u_v \cap (v_0 \cup v_1) \cup u_0 \setminus v_1 \cup u_1 \setminus v_0 \quad (3.35)$$

$$V_1 := v_0 \cup v_1 \quad (3.36)$$

erweitert wird. Die erste Aufnahme selbst wird ungeplant mit vorgegebenen Standard-Parametern akquiriert, die idealerweise ein möglichst großes Sichtfeld des Sensors ergeben.

3.5.2 Externe Rekonstruktion

3.5.2.1 Nicht zu rekonstruierende Oberflächen

Im Falle der externen Rekonstruktion existieren grundsätzlich zusätzliche Oberflächen, die zur Fixiervorrichtung des zu vermessenden Objekts gehören. Im einfachsten und am häufigsten auftretenden Fall ist das die Standfläche. In vielen Fällen wird zwar die Akquisition von Meßpunkten auf dieser Oberfläche systematisch verhindert (matt-schwarzes Material, volumetrischer Schnitt mit dem eingesehenen Volumen von Leeraufnahmen), dennoch können durch sie Verdeckungen verursacht werden. Die entsprechenden Teile der Objektrepräsentation können getrennt von der Oberfläche des

eingesehenen Volumens behandelt werden; einer zusätzlichen Vermessung soll bei der Aufnahmeplanung kein Qualitätsgewinn zugeordnet werden. Die Repräsentation erfolgt daher durch ein Dreiecksnetz, zu dem als Qualitätsattribut für alle Geometrie-elemente bereits die maximale gewünschte Qualität gespeichert wird.

3.5.2.2 Abschätzung des zu betrachtenden Volumens

Weiterhin stellt sich die Frage, wie das zu betrachtende Volumen automatisch bestimmt werden kann. Da bei der externen Rekonstruktion typischerweise reproduzierbare Beleuchtungsverhältnisse gegeben sind, bietet es sich an, hierzu eine vorgeschaltete Grobrekonstruktion des zu vermessenden Objekts auf Basis des Silhouettenschnittverfahrens unter Verwendung vordefinierter Standardpositionen einzusetzen. Hierzu wird zunächst für jede Position eine einfache Intensitäts-Leeraufnahme gespeichert, die durch Differenzbildung mit einem Intensitätsbild des Objekts dessen Silhouette liefert.

Die entstehenden Binärbilder, die für jeden Bildpunkt die Information „Objekt“ oder „Hintergrund“ tragen, können konsistent in die Oberflächendefinition integriert werden, indem für jedes Bild b_i analog zur Distanzfunktion g_i in Gleichung 3.5 eine entsprechende Funktion $\gamma_i : \mathbb{R}^3 \rightarrow \mathbb{R}$ mit

$$\gamma_i(\mathbf{x}) := \begin{cases} +1 & \text{wenn } b_i\left(\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} h(\mathbf{x})\right) = \text{„Hintergrund“} \\ -1 & \text{sonst} \end{cases} \quad (3.37)$$

eingeführt wird, die sich in ihrer Behandlung bei der Definition der Funktion f (Gleichung 3.4) nicht von den Funktionen g_i unterscheidet.

Im Gegensatz zur Datengewinnung durch den Tiefenbild-Sensor zeichnet sich das hier verwendete Differenzbildverfahren allerdings durch eine sehr viel geringere Zuverlässigkeit aus, da ein Bereich des Objekts, der im Bild die gleiche (oder geringer als ein Schwellwert abweichende) Intensität wie der von ihm verdeckte Hintergrund liefert, fehlerhaft als „Hintergrund“ klassifiziert werden kann. Die Auswirkungen einer solchen Fehlklasifikation sind durch die Maximum-Bildung in Gleichung 3.4 nicht mehr korrigierbar, so daß der Datensatz global inkonsistent und insbesondere für die Erreichbarkeitsbestimmung des Roboters unbrauchbar wird. Daher kann dieses Verfahren nicht vollautomatisch, sondern nur mit interaktiver Kontrolle und Korrektur durch den Benutzer eingesetzt werden. Aus dem gleichen Grund ist es sinnvoll, eine Dilatation der detektierten Objektregion im Bild um ein bis zwei Bildpunkte vorzunehmen.

3.6 3D-Rekonstruktion aus mehreren Einzelbildern

Vergleichend zur inkrementellen Konstruktion einer Repräsentation der vermessenen Oberfläche soll in diesem Abschnitt die Vorgehensweise am Ende des Akquisitionsprozesses skizziert werden, mit der aus einer Anzahl überlappender, registrierter Einzelaufnahmen ein Modell des vermessenen Objekts generiert wird. Die Kenntnis der späteren Verwendung der akquirierten Meßdaten bildet eine wichtige Voraussetzung für die Bewertung des Nutzens einer zusätzlichen Aufnahme im Planungsprozess. Dabei wird auf die Umsetzung im SCULPTOR-System Bezug genommen, das die Grundlage eines Teiles der Ergebnisse der vorliegenden Arbeit bildet [72].

Anders als in der vorliegenden Arbeit wird im SCULPTOR-System nicht die inkrementelle Ergänzung des Datensatzes, sondern die globale Mehrbildintegration als zentrale Strategie verfolgt. Als Grundvoraussetzung müssen hierzu alle Einzelaufnahmen näherungsweise in ein gemeinsames Koordinatensystem eingepaßt worden sein. Die Näherungslösungen für die extrinsischen Parameter jeder Aufnahme können entweder durch das Positioniersystem oder durch Punktkorrespondenzen gewonnen werden. Um eine Akkumulation der dabei auftretenden Restfehler zu vermeiden, werden die extrinsischen Parameter des Datensatzes zuerst für den gesamten Bildverbund simultan optimiert. In den letzten Jahren wurden diverse Verfahren zur Lageoptimierung großer Tiefenbildverbünde vorgeschlagen [71, 83], so daß die Lösung dieser Aufgabe als Stand der Technik anzusehen ist. Im Folgenden werden nun die weiteren Schritte zur Rekonstruktion der vermessenen Oberfläche aus dem bereits registrierten Mehrbilddatensatz beschrieben.

3.6.1 Redundanznutzung in mehrfach vermessenen Bereichen

Nach Abschluß der Bildakquisition sind typischerweise wesentliche Teile der zu rekonstruierenden Oberfläche in mehr als einer Aufnahme sichtbar, so daß sich die Frage nach der Fusion der redundanten Meßdaten stellt. Im Rahmen der Fusion sollen mehrere Ziele erreicht werden:

- Das Rauschen der Meßdaten soll durch Mittelwertbildung reduziert werden.
- Meßfehler sollen durch Konsistenzprüfung erkannt und eliminiert werden.
- Bildbereiche, die einen mehrfach vermessenen Oberflächenabschnitt mit vergleichsweise niedriger Auflösung repräsentieren, sollen bei der interpolierenden Berechnung von Oberflächenpunkten nicht berücksichtigt werden.

Diese Ziele lassen sich auf dem bereits registrierten Bildverbund durch Analyse der bei den Einzelmessungen, also bildpunktweise, vorliegenden Blickrichtungen des Sensors erreichen. Hierzu wird der Sichtstrahl des Sensors c zu einem gegebenen Meßpunkt p betrachtet, so daß alle in der gegebenen Pixelposition von p speicherbaren

Meßpunkt-Positionen als Punkte auf einer parametrisch beschriebenen Gerade durch

$$\mathbf{p} + \lambda \mathbf{c} \quad (3.38)$$

dargestellt werden können. Die zu berücksichtigenden Meßpunkte in anderen Aufnahmen können durch Schnittpunktbestimmung dieser Geraden mit der vermessenen Oberfläche in einer geeignet gewählten Umgebung von \mathbf{p} bestimmt werden, so daß für n zu berücksichtigende Meßpunkte letztlich eine Menge von skalaren Geradenparametern $\lambda_1, \dots, \lambda_n$ vorliegt. Bei der Kombination dieser Einzelwerte werden drei Gewichte verwendet: die Qualität w_q der Datenpunkte (Abtastdichte sowie weitere sich aus dem Meßverfahren ergebende Maßzahlen), ein durch das *M-estimates*-Verfahren (Tukey's biweight function) gewonnenes Gewicht w_t , welches Ausreißer eliminiert, sowie ein Distanzmaß w_d , welches den Abstand zum nächsten Rand des kontinuierlich vermessenen Oberflächenbereichs im jeweiligen Bild angibt. Durch Anwendung dieser Gewichte kann nun eine gewichtete Mittelwertbildung

$$\lambda = \sum_i w_q^{(i)} \cdot w_t^{(i)} \cdot w_d^{(i)} \cdot \lambda_i \quad (3.39)$$

vorgenommen werden, so daß Meßfehler eliminiert werden und ein glatter Übergang an den Rändern der kontinuierlich vermessenen Bereiche der Bilder erreicht wird. Zusätzlich werden Bildbereiche, die die Meßdaten durch die Rasterauflösung des Bildes an der gemessenen 3D-Position nur in geringer Auflösung repräsentieren können, ausmaskiert, sofern die jeweiligen Meßwerte in anderen Bildern in höherer Auflösung repräsentiert werden können.

Die redundante Mehrfachvermessung von Oberflächenbereichen erhöht also nicht nur den Datenumfang, sondern trägt auch zur Qualitätsverbesserung bei. Das oben skizzierte Verfahren ist jedoch nicht in der Lage, aus den redundanten Punkterastern eine kombinierte Oberflächenbeschreibung mit einer höheren lateralen Oberflächenauflösung als in einem der einfließenden Bildraaster zu gewinnen, so daß zum Erreichen einer vorgegebenen Abtastdichte dennoch ein Bild erforderlich ist, das den betreffenden Oberflächenbereich mit mindestens dieser Abtastdichte zeigt.

3.6.2 Oberflächendefinition und Triangulierung

Die Oberflächendefinition basiert im SCULPTOR-System auf einer impliziten Definition ähnlich der in Abschnitt 3.3 durch die Gleichungen 3.3 und 3.4 definierten Funktion. Da die Tiefenbilder nicht inkrementell, sondern erst nach Akquisition aller Aufnahmen weiterverarbeitet werden, wird hier eine andere Ergänzung des Definitionsbereichs vorgenommen, indem statt Gleichung 3.4 folgende Definition verwendet wird:

$$f(\mathbf{x}) := \begin{cases} \max_i g_i(\mathbf{x}) & \exists i \mid r_i(u, v) \text{ ist definiert} \\ +1 & \text{sonst} \end{cases} \quad (3.40)$$

Dabei ist wieder (u, v) die Pixelposition in Bild i , auf die \mathbf{x} projiziert wird. Auf diese Weise wird das betrachtete Volumen, in dem Nulldurchgänge von f auftreten können, eingeschränkt, ohne daß hierzu eine Grenze des zu rekonstruierenden Bereichs angegeben werden muß. Der Grundgedanke dieses Ansatzes ist die Annahme, daß das Objekt bereits vollständig vermessen wurde, so daß sich in den Volumenbereichen, die nicht auf definierte Tiefenbildoberflächen projiziert werden können, kein Teil des Objektes befinden kann.

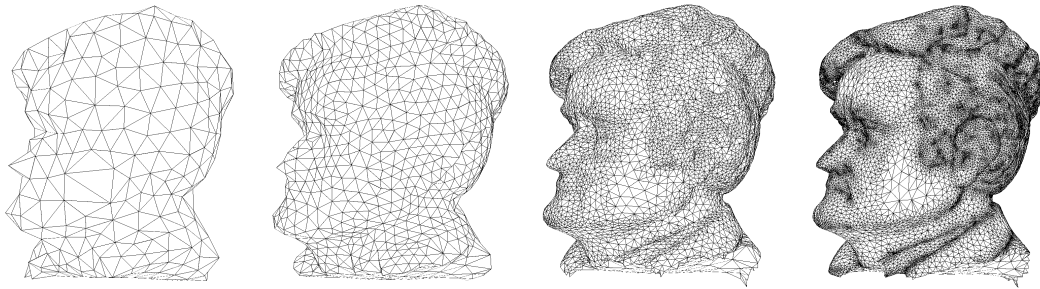


Abbildung 3.11: Sukzessive Verfeinerung einer groben Initialtriangulierung durch adaptive Dreiecksunterteilung

Für die weitere Verarbeitung in der Computergraphik und im *rapid prototyping* ist die Repräsentation der definierten Objektoberfläche durch ein Dreiecksnetz die Methode der Wahl. Zur Triangulierung implizit definierter Oberflächen wurde eine große Zahl von Verfahren vorgeschlagen [61, 69, 10], so daß die Lösung dieses Problems ebenfalls als Stand der Technik betrachtet werden kann. Es sei noch darauf hingewiesen, daß die Anzahl der akquirierten Einzelmeßpunkte im Allgemeinen abgesehen vom Speicherplatzbedarf keinen direkten Einfluß auf die Laufzeit der in Frage kommenden Verfahren hat, da die Oberfläche zunächst durch eine sehr weitmaschige Initialtriangulierung repräsentiert werden kann, die nachfolgend durch sukzessive Verfeinerung an die jeweilige lokale Oberflächenkrümmung angepaßt wird [73] (Abbildung 3.11).

3.7 Zusammenfassung eigener Ergebnisse

In diesem Kapitel wurde zunächst ein Ansatz zur Repräsentation der teilvermessenen Objektoberfläche vorgeschlagen, der auch die unvermessenen Bereiche sinnvoll repräsentiert. Dabei bildet der Begriff des „eingesehenen Volumens“ die zentrale Überlegung, die zwei disjunkte Teilräume unterscheidet: den eingesehenen und daher bekanntermaßen leeren Teilraum, sowie den Teilraum, der die zu vermessende Objektoberfläche enthält. Die so definierte Oberfläche wird zusammen mit qualitätsbezogenen Oberflächenattributen im nachfolgenden Kapitel als Grundlage einer Bewertungs-

funktion potentieller Aufnahmekonfigurationen eingesetzt, so daß die Effizienz der Repräsentationsform und des dazugehörigen Konstruktionsverfahrens entscheidend für das Gesamtverfahren ist.

Im Einzelnen wurde zunächst die formale Definition einer vorzeichenbehafteten Volumenfunktion hergeleitet, die zwischen der Situation bei der Rekonstruktion von Innenräumen (interne Rekonstruktion, das Meßsystem ist von der zu rekonstruierenden Oberfläche eingeschlossen) und von Gegenständen (externe Rekonstruktion, das Meßvolumen ist innerhalb des Arbeitsbereichs unbegrenzt und das Meßsystem befindet sich außerhalb der zu rekonstruierenden Oberfläche) unterscheidet.

Der wesentliche Beitrag des Kapitels besteht in den vorgeschlagenen Verfahren zur inkrementellen Konstruktion der so definierten Oberfläche und der Berechnung, Repräsentation und computergraphischen Darstellung von qualitätsbezogenen Maßzahlen, die jedem Punkt der Oberfläche als Attribut zugeordnet werden. Hier wird unterschieden zwischen dem vermessenen Teil der Oberfläche, der durch eine Vereinigungsoperation inkrementell durch zusätzliche Aufnahmen ergänzt wird, sowie der unvermessenen Grenzfläche des eingesehenen Volumens, die durch eine Schnittoperation bezüglich des eingesehenen Volumens inkrementell durch zusätzliche Aufnahmen reduziert wird.

Die vorgeschlagenen Verfahren orientieren sich an der beabsichtigten Berechnung einer Bewertungsfunktion mit Hilfe computergraphischer Verfahren. Daher wird die inkrementelle Konstruktion der Oberflächenelemente nicht, wie aus der Literatur bekannt, auf Basis von Volumendaten durchgeführt, sondern direkt auf den computergraphischen Primitiven, so daß diese soweit möglich nach der Integration neuer Aufnahmen weiterverwendet werden können. Die redundante Durchführung von Vereinfachungsschritten auf dem Dreiecksnetz kann so vermieden werden.

Mit Hilfe der vorgeschlagenen Verfahren wird eine computergraphisch darstellbare Geometrie erzeugt, mit der die Sichtbarkeit eines Punktes von den Beobachtungspunkten des Meßsystems aus entschieden, die bereits erreichte Meßdatenqualität abgelesen, und die zu erwartende Datenqualität vorberechnet werden kann. Damit sind die Grundlagen für die im folgenden Kapitel vorgeschlagenen Verfahren zur Bewertung und Optimierung von Aufnahmekonfigurationen gelegt.

KAPITEL 4

Ziel und Umsetzung der Aufnahmeplanung

4.1 Einleitung

Aufbauend auf der im vorangegangenen Kapitel vorgeschlagenen Konstruktion einer Oberflächenbeschreibung, die auch unvermessene Teile des Objekts repräsentiert, soll in diesem Kapitel nun ein Verfahren entwickelt werden, welches die Sequenz der Aufnahmekonfigurationen dergestalt bestimmt, daß das Objekt mit möglichst geringem Aufwand möglichst vollständig mit einer vorgegebenen Abtastdichte erfaßt wird. Die Lösung dieser Aufgabe kann in mehrere Teilprobleme mit unterschiedlichem Abstraktionsgrad gegliedert werden:

- Gegebene Aufnahmekonfigurationen müssen anhand der jeweils vorliegenden Repräsentation des Objekts durch eine geeignete Metrik bewertet werden können, so daß die Aufnahmesequenz iterativ nach dem *greedy*-Ansatz bestimmt werden kann.
- Neben dem Aufwand der Umsetzung geplanter Aufnahmekonfigurationen muß auch der Rechenaufwand der Planung selbst berücksichtigt werden. Daher wird eine Hierarchie mehrerer Metriken benötigt, die mit zunehmendem Aufwand und entsprechend zunehmender Ergebnisqualität gegebene Aufnahmekonfigurationen bewerten, so daß ein mehrstufiges Verfahren zur Bestimmung der Aufnahmesequenz aufgebaut werden kann.
- Zur Reduktion des Rechenaufwands ist es wünschenswert, die Ergebnisse der Bewertung in Frage kommender Aufnahmekonfigurationen zumindest teilweise zu speichern und nach der Akquisition neuer Aufnahmen inkrementell zu aktualisieren, so daß redundante Evaluationen der Metrik vermieden werden können.

Die Umsetzung von Aufnahmekonfigurationen und die weitere Verarbeitung der erfaßten Daten, insbesondere ihre Registrierung, implizieren außerdem Nebenbedingungen, die bei der Bestimmung der Aufnahmekonfigurationen berücksichtigt werden müssen. Dabei treten konkurrierende Ziele auf: während einerseits die Erfassung möglichst großer, bis jetzt unvermessener Bereiche der Objektoberfläche angestrebt wird, erschwert genau dies andererseits die zuverlässige Registrierung dieser Oberflächenbereiche mit den bis dahin vorliegenden Meßdaten.

Neben diesen Problemstellungen, die sich auf die Bestimmung erwünschter Aufnahmekonfigurationen beziehen, treten darüberhinaus auch Situationen auf, die eine Vermeidung bestimmter Aufnahmekonfigurationen sinnvoll erscheinen lassen, ohne daß dies aus der Repräsentation des Objekts, wie sie in Kapitel 3 vorgeschlagen wurde, allein abgeleitet werden könnte. Stattdessen erfordern diese im folgenden genannten Situationen die Berücksichtigung der Aufnahmesequenz selbst, also bereits erfolgter oder zukünftig möglicher Aufnahmen:

- Aufnahmekonfigurationen, die bereits zu einem erfolglosen Versuch einer Datenakquisition geführt haben, sollten nicht wiederholt werden. Dies ist aus der Objektrepräsentation allein zunächst nicht ersichtlich, da der betroffene unvermeßbare Teil der Oberfläche seinen unvermessenen Zustand nicht ändert.
- Kleine und mittelgroße Objekte werden typischerweise im Verlauf ihrer Digitalisierung bewegt, um beispielsweise auch die Standfläche zu erfassen. Suboptimale Aufnahmen von schwer einsehbaren Bereichen der Oberfläche sollten daher vermieden werden, wenn diese Bereiche zu einem späteren Zeitpunkt in höherer Qualität erfaßt werden können.

Weiterhin können drei verschiedene Abschnitte mit jeweils unterschiedlichen Zielsetzungen innerhalb der zu planenden Aufnahmesequenz unterschieden werden (vgl. [39]):

1. Zuerst muß die initial unbekannte Geometrie soweit erfaßt werden, daß neue Aufnahmekonfigurationen sinnvoll bestimmt und Kollisionen bei ihrer Umsetzung ausgeschlossen werden können.
2. Anschließend werden verbleibende Lücken im bereits vermessenen Teil der Oberfläche gezielt geschlossen.
3. Der letzte Abschnitt hat das Ziel, die vorgegebene Abtastdichte zu erreichen.

Diese Abschnitte können auch als unterschiedliche Fälle der Aufnahmeplanung gesehen werden, bei denen die zu vermessende Geometrie initial bereits in grober Näherung beziehungsweise mit hoher Genauigkeit vorliegt, so daß der erste beziehungsweise der die ersten beiden Abschnitte nicht ausgeführt werden müssen.

Aus dieser kurzen Beschreibung der einzelnen Aspekte wird bereits deutlich, daß es sich bei der Aufnahmeplanung um eine relativ komplexe Problemstellung mit mehreren, teilweise konkurrierenden Zielen handelt. In den folgenden Abschnitten dieses

Kapitels soll, nach einem Überblick zum Stand der Technik, zunächst eine theoretische Grundlage zur Beschreibung der oben genannten Aspekte gelegt werden. Anschließend werden Vorschläge zur Bestimmung einer Lösung gemacht; dabei liegt das Hauptaugenmerk auf einer effizienten Berechnung in vertretbarem Zeitaufwand, damit das vorgeschlagene Verfahren auch praktisch genutzt werden kann. Neben der grundsätzlichen Struktur des Planungsalgorithmus bildet daher das Verfahren zur Berechnung der Zielfunktion einen der Schwerpunkte des Kapitels. Ein Vorschlag für ein geeignetes Abbruchkriterium des iterativen Gesamtverfahrens bildet den Abschluß des Kapitels.

4.2 Stand der Technik

Die planungstheoretische Betrachtung der vorliegenden, aus der Literatur als *“next best view”*-Problem bekannten Aufgabenstellung wird in den veröffentlichten Arbeiten zur Aufnahmeplanung meist stiefmütterlich behandelt. Monographien zur Planung [55, 57] behandeln die Aufnahmeplanung auf der anderen Seite als eine von vielen relevanten Problemstellungen, denen nur wenig Raum zur detaillierten Betrachtung gegeben wird. Daher werden in der vorliegenden Arbeit die relevanten Datenstrukturen, Planungs- und Akquisitions-Schritte im Kontext der in den genannten Monographien beschriebenen Konzepte dargestellt.

In den bisher veröffentlichten Arbeiten zur Aufnahmeplanung wird typischerweise nur ein niedrig-dimensionaler Suchraum unterstützt, so daß eine geeignet gewählte Zielfunktion einfach durch uniforme Abtastung des Parameterraums optimiert werden kann. Pito [80] schlägt hierzu als zusätzliche Datenstruktur den sogenannten *“positional space”* vor, der die Planungsaufgabe unter Wiederverwertung bekannter, unveränderter Ergebnisse aus der vorigen Iteration unterstützt. Der Nachteil dieser Ansätze besteht darin, daß sie nicht auf höherdimensionale Parameterräume übertragen werden können. In der Literatur konnten keine Ansätze zur Aufnahmeplanung gefunden werden, die Parameterräume der Dimensionalität 6 oder höher unter Berücksichtigung einer inhomogenen räumlichen Verteilung von Nebenbedingungen abtasten.

In anderem Kontext wurden zu diesem Problem allerdings seit langem Lösungsvorschläge beigetragen. In der numerischen Integration werden Punktfolgen benötigt, die eine räumlich gleichmäßige Abtastung eines mehrdimensionalen Parameterraums ermöglichen. Hierzu sind neben der in der vorliegenden Arbeit eingesetzten *van-der-Corput*-Punktfolge [20] vor allem die gebräuchlicheren *Halton*- [40] und *Hammersley*-Punktfolgen [41] zu nennen (alle zitiert nach [57]). Niederreiter [76] nennt weitere alternative Ansätze, die allerdings im Rahmen der vorliegenden Arbeit nicht näher untersucht wurden.

Weitere relevante Vorarbeiten werden des einfacheren Verständnis' wegen an der Stelle der Beschreibung der jeweiligen neuen Beiträge genannt. Stellvertretend sei hier noch auf die Übersichtsarbeiten von Durand zur Sichtbarkeit [29] und von Hasenfratz et al. zur Schattenrechnung [44] verwiesen.

4.3 Formulierung des Problems

In diesem Abschnitt soll das Problem der Aufnahmeplanung für initial unbekannte Geometrien formal beschrieben und in Beziehung zur Planungstheorie gesetzt werden. In den einzelnen Unterabschnitten werden dabei neben der formalen Zielsetzung die Strukturierung des Suchraums, die einzuhaltenden Nebenbedingungen sowie der zu erwartende Rechenaufwand und die sich daraus ergebende Notwendigkeit einer vereinfachten Näherung thematisiert.

4.3.1 Zielsetzung

Die konkrete Zielsetzung der Aufnahmeplanung für initial unbekannte Objekte unter Berücksichtigung des Aufwands läßt sich folgendermaßen formulieren:

Finde diejenige Menge von Aufnahmekonfigurationen, mit der die Vereinigung der erfaßten Meßdaten eine vollständige Abdeckung der Objektoberfläche mit einer vorgegebenen Abtastdichte ergibt, und die eine gegebene Kostenfunktion minimiert.

Für die weitere Betrachtung ist es hilfreich, die Eigenschaften der einzelnen Elemente dieser Formulierung näher zu beschreiben und auf einen planungstheoretischen Kontext zu beziehen. Dabei wird auf die von LaValle [57] vorgeschlagenen Konzepte Bezug genommen.

Der Planungsvorgang läuft in diskreten Zeitschritten $k \in \mathbb{N}$ ab, zwischen denen lediglich Verarbeitungsschritte durchgeführt werden, die keiner Planung bedürfen. Die Bewegung des Roboters, die eigentliche Bildakquisition und die Integration des Bildes in den bestehenden Bildverbund (einschließlich der in Kapitel 3 beschriebenen Hilfsstrukturen) bleiben daher an dieser Stelle unberücksichtigt.

Zu jedem Zeitschritt k befinden sich Aufnahmesystem und Arbeitsraum in einem bestimmten Zustand x_k aus der Menge möglicher Zustände X . Der Zustand des Arbeitsraums $e \in E$ kann dabei als konstant angenommen werden; andernfalls wäre die Rekonstruktion des darin befindlichen Objekts durch Mehrbildverfahren im Allgemeinen unmöglich. Die Konfiguration des Aufnahmesystems, konkret seine Lage relativ

zur Objektoberfläche, kann durch Registrierung mit hoher Genauigkeit bestimmt werden und wird daher als bekannt angenommen. Weitere Eigenschaften des Zustandes von Aufnahmesystem und Arbeitsvolumen sind für die Planungsaufgabe nicht relevant. Formal kann jeder Zustand $x_k \in X$ daher in eine Konfiguration q_k des Aufnahmesystems und den konstanten Zustand der Umgebung e mit $x_k = (q_k, e)$ zerlegt werden.

Dem Planungsalgorithmus stehen Informationen über das Arbeitsvolumen allerdings nur eingeschränkt in Form der erfaßten und verarbeiteten Daten zur Verfügung. Der Umfang dieser Informationen wird durch jeden neuen Akquisitionsschritt ergänzt. Anstelle des beispielsweise in der Bewegungsplanung für Roboter klassischerweise anzutreffenden Zustandsraums X basiert die Beschreibung des Planungsproblems daher auf dem Informationsraum I , aus dem für jeden Zeitschritt ein Informationsstand i_k vorliegt.

Durch die Betrachtung der in Kapitel 3 eingeführten Objektrepräsentation wird deutlich, daß die Abgrenzung zwischen zustandsbasierter und informationsstandbasierter Planung im vorliegenden Fall relativ unscharf ist. Bei der zustandsbasierten Planung wird zunächst der Zustand aus der vorliegenden Information geschätzt. Da die vorliegende Information zur Objektgeometrie und zum Arbeitsvolumen unvollständig ist, muß ein zustandsbasierter Planungsalgorithmus robust gegenüber fehlerhaft geschätzten Zuständen sein. Alternativ wird die Planung bei der informationsstandbasierten Planung direkt auf Basis des Informationsstands durchgeführt, so daß der tatsächliche oder geschätzte Zustand für die Planung nicht benötigt wird.

Das in der vorliegenden Arbeit vorgeschlagene Verfahren nutzt keine expliziten Methoden zur Schätzung des Zustands, sondern basiert auf einer geometrischen Modellierung des Informationsstandes. Der Informationsstand bezieht sich damit nicht nur auf den aktuellen Zustand, sondern auch auf alle vorangegangenen Zustände des Systems während der bisherigen Ausführung des Planungsalgorithmus sowie auf die dabei vom Sensor erfaßten Daten. In der Literatur wird ein derartiger Informationsraum als "history information space" [57] bezeichnet.

Formal ist der Plan

$$\pi : I \rightarrow U \quad (4.1)$$

damit in jedem Schritt eine Funktion des Informationsstandes i_k und berechnet aus diesem eine auszuführende Aktion $u_k \in U$ aus der Menge der möglichen Aktionen. Ziel des Planungsalgorithmus ist die Berechnung einer Sequenz $\{u_1, \dots, u_n\}$ von Aktionen, welche zu einer konstanten Umgebung $e \in E$ eine Menge von Meßpunkten auf der Oberfläche liefert, deren paarweiser Abstand einen vorgegebenen Schwellwert nicht überschreitet (vorgegebene Mindestabtastdichte). Als konkurrierendes Ziel soll zusätzlich der Aufwand des Plans minimiert werden. Mit einer Kostenfunktion

$c : U \rightarrow \mathbb{R}$ können die Gesamtkosten durch

$$\sum_{k=1}^n c(u_k) \quad (4.2)$$

ausgedrückt werden. Dabei ist die tatsächlich benötigte Anzahl von Einzelnahmen n initial unbekannt.

In der nachfolgend beschriebenen Problemstellung setzen sich die konkret repräsentierten Elemente des Informationsraums jeweils aus der Konfiguration des Aufnahmesystems in allen vorangegangenen Schritten sowie der in Gleichung 3.7 definierten Oberfläche des bisher eingesehenen Volumens mit

$$I_k = \mathbb{S}^2 \times \tilde{U}_k \quad (4.3)$$

und (entsprechend [57])

$$\tilde{U}_k = \underbrace{U \times \cdots \times U}_k \quad (4.4)$$

zu

$$i_k = (\mathbf{S}_k, u_1, \dots, u_k) \quad (4.5)$$

zusammen, dabei beschreibt \mathbb{S}^2 die Menge der in den dreidimensionalen Raum eingebetteten 2-Mannigfaltigkeiten (Oberflächen physikalisch realisierbarer Objekte).

Zur Vereinfachung der Schreibweise wird hierbei die geplante Aktion u_i und die tatsächlich ausgeführte Aktion $u_i + \Delta u_i$ synonym verwendet, da die verwendeten Robotersysteme eine hohe Genauigkeit bei der Umsetzung der jeweiligen Aktion erreichen, so daß Δu_i für die Planung vernachlässigt werden kann. In der inkrementellen Konstruktion der Objektoberfläche wird die durch den Fehler der Aktionsausführung Δu_i verursachte Lageungenauigkeit bei der Integration des zusätzlichen Bildes durch Registrierung mit hoher Genauigkeit bestimmt und rechnerisch kompensiert. Entsprechend enthält i_{k+1} nach Gleichung 4.2 nicht direkt das Ergebnis $\pi(i_k)$ des vorangegangenen Planungsschrittes, sondern die Daten der tatsächlich ausgeführten Aktion, so daß Gleichung 4.5 bei Berücksichtigung des Fehlers der Aktionsausführung

$$i_k = (\mathbf{S}_k, u_1 + \Delta u_1, \dots, u_k + \Delta u_k) \quad (4.6)$$

lautet.

Bevor in Abschnitt 4.4 gezeigt wird, wie auf Basis dieses Informationsstandes die Planung durchgeführt werden kann, werden im Folgenden zunächst die Eigenschaften des Aktionsraums U beschrieben.

4.3.2 Aktions- und Suchraum

Das Ergebnis jedes Planungsschrittes ist ein vollständiger Satz von Parametern zur Beschreibung der nächsten Bildakquisition. Abhängig von den zur Laufzeit variablen Größen des Systems umfaßt der Aktionsraum U daher im Allgemeinen extrinsische und intrinsische Parameter des Sensors. Analog zur Benennung in der Photogrammetrie bezeichnen die extrinsischen Parameter die Lage des Sensors im Raum, die vom Roboter umgesetzt werden muß, während sich die intrinsischen Parameter auf die Bildgeometrie des Sensors beziehen.

4.3.2.1 Extrinsische Parameter

Konkret wurden die in Frage kommenden extrinsischen Parameter in Kapitel 2.4 beschrieben und insbesondere die Einschränkung der Schnittstelle auf eine Orientierung gegeben durch sechs Freiheitsgrade erwähnt. Für Zwecke der Planung ist es jedoch sinnvoll, den entsprechenden Aktionsraum (drei Drehwinkel und drei Translationsparameter) auch in Bezug auf seine topologische Struktur zu analysieren. Dabei ist insbesondere die Parametrisierung der Menge aller Rotationen im dreidimensionalen Raum aufgrund der Singularitäten und Identitäten einer einfachen Winkelrepräsentation problematisch: die gleiche Rotationsmatrix kann beispielsweise durch unterschiedliche Kombinationen von Eulerwinkeln ausgedrückt werden.

In der Mathematik der Bewegungsplanung für Roboter (hier dargestellt nach Latombe [55]) wird die Menge aller n -dimensionalen Rotationsmatrizen als *spezielle orthogonale Gruppe* $SO(n)$ bezeichnet. Mit dieser Gruppe der Rotationen kann die Menge aller $(n + 1) \times (n + 1)$ -Transformationsmatrizen, die eine Rotation und Translation beschreiben, definiert werden. Diese haben die Form

$$\left\{ \begin{pmatrix} R & v \\ 0 & 1 \end{pmatrix} \mid \text{mit } R \in SO(n) \text{ und } v \in \mathbb{R}^n \right\} \quad (4.7)$$

und werden als *spezielle euklidische Gruppe* $SE(n)$ bezeichnet. Da Rotation und Translation unabhängig voneinander sind ist die Gruppe $SE(n)$ homöomorph¹ zur Gruppe $SO(n) \times \mathbb{R}^n$.

Die Problematik der Parametrisierung dieser Gruppen soll aus Gründen der Anschaulichkeit zunächst am einfachen Beispiel der Gruppe $SE(2)$, also der *rigid body*-Transformationen im zweidimensionalen Raum, betrachtet werden. Hier bildet die Translationskomponente $(t_x, t_y) \in \mathbb{R}^2$ eine zweidimensionale Mannigfaltigkeit. Die unabhängig wählbare Rotation $\vartheta \in [0, 2\pi)$ bildet wegen der Identität durch die Äquivalenz $0 \sim 2\pi$ eine eindimensionale Mannigfaltigkeit \mathbb{S}^1 . Daher ist $SE(2)$ homöomorph

¹Zwei topologische Räume werden als homöomorph bezeichnet, wenn eine bijektive und kontinuierliche Abbildung zwischen ihnen existiert.

zu $\mathbb{S}^1 \times \mathbb{R}^2$; man kann sich dies als eine in den vierdimensionalen Raum eingebettete dreidimensionale Zylinderoberfläche vorstellen; eine Parametrisierung von $SE(2)$ muß diese Topologie entsprechend berücksichtigen.

Die entsprechende Topologie für $SO(3)$ und damit auch für $SE(3)$ ist sehr viel schwieriger zu beschreiben als im zweidimensionalen Fall. Hierzu wird üblicherweise die Quaternionen-Repräsentation der Rotation als Hilfsstruktur verwendet. Ein Quaternion $q = (s, v) \in \mathbb{R}^4$ besteht ähnlich den komplexen Zahlen aus einem Realteil $s \in \mathbb{R}$ und einem dreidimensionalen Imaginärteil $v \in \mathbb{R}^3$, welche als skalarer Drehwinkel und vektorielle Drehachse interpretiert werden können und so eine dreidimensionale Rotation beschreiben. Mit der Einschränkung auf Einheitsquaternionen mit $\|q\| = 1$ kann einem Quaternion $q = (s, (v_x, v_y, v_z)^T)$ eindeutig eine Rotationsmatrix durch

$$R(q) = \begin{pmatrix} 2(s^2 + v_x^2) - 1 & 2(v_x v_y - s v_z) & 2(v_x v_z + s v_y) \\ 2(v_x v_y + s v_z) & 2(s^2 + v_y^2) - 1 & 2(v_y v_z - s v_x) \\ 2(v_x v_z - s v_y) & 2(v_y v_z + s v_x) & 2(s^2 + v_z^2) - 1 \end{pmatrix} \quad (4.8)$$

zugeordnet werden. Umgekehrt kann mit den aus einer Rotationsmatrix gewonnenen Drehachse v mit $\|v\| = 1$ und Drehwinkel ϑ ein Einheitsquaternion durch

$$q = \pm \left(\cos \frac{\vartheta}{2}, \sin \frac{\vartheta}{2} \cdot v \right) \quad (4.9)$$

gebildet werden; diese Zuordnung ist wegen $R(q) = R(-q)$ jedoch nicht eindeutig. Algebraisch bilden die Einheitsquaternionen mit ihrer Verkettungsoperation

$$q_1 \cdot q_2 = (s_1 s_2 - v_1 v - 2, s_1 v_2 + s_2 v_1 + v_1 \times v_2), \quad (4.10)$$

der Inversenbildung $q^{-1} = (s, -v)$ und dem neutralen Element $e = (1, (0, 0, 0)^T)$ (ebenso wie $SO(3)$) eine Gruppe², die üblicherweise als *spezielle Gruppe der Einheitsquaternionen*³ $SU(2)$ bezeichnet wird. Die in Gleichung 4.8 definierte Zuordnungsfunktion $R : SU(2) \rightarrow SO(3)$ bildet dabei einen Gruppenhomomorphismus, da die Beziehungen $R(q_1 \cdot q_2) = R(q_1) \cdot R(q_2)$ sowie $R(e) = I$ erfüllt sind.

²Eine Gruppe ist eine Menge G , die zusammen mit einer binären Operation \cdot die folgenden Gruppenaxiome erfüllt:

Abgeschlossenheit: Für alle $a, b \in G$ gilt $a \cdot b \in G$

Assoziativität: Für alle $a, b, c \in G$ gilt $(a \cdot b) \cdot c = a \cdot (b \cdot c)$

Neutrales Element: Es gibt ein $e \in G$ mit $e \cdot a = a \cdot e = a$ für alle $a \in G$

Inverse: Für jedes $a \in G$ existiert ein $a^{-1} \in G$ mit $a \cdot a^{-1} = a^{-1} \cdot a = e$

³Die Benennung ist auf den englischen Begriff *unit quaternion* zurückzuführen.

Topologisch bildet die Menge der Einheitsquaternionen wegen der Bedingung $\|q\| = 1$ für alle $q \in SU(2)$ eine in den vierdimensionalen Raum eingebettete dreidimensionale Oberfläche der Einheitskugel, also eine 3-Mannigfaltigkeit, die mit \mathbb{S}^3 bezeichnet wird. Aufgrund der in Gleichung 4.9 gegebenen zweideutigen Zuordnung zu $SO(3)$ ist \mathbb{S}^3 aber nicht homöomorph zu $SO(3)$. Mit der Identifikation durch die Äquivalenz $q \sim -q$ für alle $q \in SU(2)$ ergibt sich topologisch vielmehr die Homöomorphie zu \mathbb{P}^3 , dem dreidimensionalen projektiven Raum. Dieser kann als die Menge aller den Ursprung enthaltenden Geraden im vierdimensionalen Raum betrachtet werden, so daß die Punkte p und $-p \in \mathbb{R}^4$ mit $\|p\| = 1$ die identische Gerade durch den Ursprung beschreiben, wie auch die entsprechenden Rotationen identisch sind. Daher ist $SO(3)$ homöomorph zu \mathbb{P}^3 und

$$SE(3) \cong \mathbb{P}^3 \times \mathbb{R}^3. \quad (4.11)$$

Die Struktur des Suchraums wird in Abschnitt 4.6 noch eine wichtige Rolle spielen.

4.3.2.2 Intrinsische Parameter

Während die extrinsischen Parameter bei der Planung für einen beweglichen Sensor sozusagen natürlicherweise Gegenstand der Optimierung sind, ist dies für die intrinsischen Parameter weniger naheliegend und tatsächlich auch nur für spezielle Sensortypen der Fall. Ursache hierfür sind meist Einschränkungen der Hardware, die beispielsweise eine Änderung der Brennweite eines perspektivischen Kamerasystems nicht zulassen, ohne daß eine Neukalibrierung des Geräts erforderlich wird.

Zur Verallgemeinerung dieses Sachverhalts kann die in Abschnitt 2.2 verwendete Unterteilung [7] in den eigentlichen Sensor, der punkt-, zeilen- oder bildweise Meßdaten erfaßt, und die zusätzlichen Ablenk- oder Positioniereinheiten herangezogen werden. Typischerweise sind diese zusätzlichen Ablenk- oder Positioniereinheiten zur Laufzeit programmierbar, so daß das sequentiell vom Sensor erfaßte Bild in seiner Abtastdichte (Schrittweite) und -bereich modifiziert werden kann.

Konkret trifft dies auf die im Rahmen der vorliegenden Arbeit eingesetzten Lichtlaufzeitsysteme zu, deren horizontale und vertikale Ablenkung durch eine frei programmierbare *pan-tilt unit* realisiert ist, so daß sowohl die Abtastdichte als auch der Bildausschnitt frei wählbar sind⁴. Im Einzelfall muß mit Blick auf den Aufbau des Meßsystems untersucht werden, welche Parameter sinnvoll und unabhängig von den extrinsischen Parametern in die Planung einbezogen werden können. Beispielsweise kann

⁴Im Unterschied zu kamerabasierten Sensoren kann dabei durch die sequentielle Bildakquisition auch tatsächlich ein Geschwindigkeitsvorteil durch die Einschränkung auf einen kleinen Bildausschnitt erzielt werden. Die Einschränkung auf eine bestimmte *region of interest* mit entsprechender Erhöhung der Bildrate ist zwar auch für viele Flächenkameras möglich; dies wird aber von marktgängigen Meßsystemen nicht unterstützt.

der horizontale Bildausschnitt für polare Sensorgeometrien auch durch entsprechende extrinsische Parameter (Rotation um die Sensorachse) definiert werden.

Damit ergeben sich für typische Lichtlaufzeitsysteme drei bis vier Parameter zur Bestimmung von Bildausschnitt und -größe, sowie bis zu zwei weitere Parameter für die horizontale und vertikale Bildauflösung. Letztere sind für marktgängige Meßsysteme allerdings meist nur in wenigen Stufen konfigurierbar. Für kamerabasierte Meßsysteme entfällt die Planung der intrinsischen Aufnahmeparameter, obwohl diese im Rahmen des in der vorliegenden Arbeit vorgeschlagenen Verfahrens grundsätzlich möglich wäre. Denkbar wäre neben dem sogenannten *binning* (Zusammenfassung von $2^n \times 2^n$ Pixeln zu einem Pixel mit entsprechend größerer Raumwinkelabdeckung) insbesondere auch eine Einschränkung auf Bildausschnitte bei der triangulationsbasierten Vermessung. Für Projektor-Kamera-Systeme ist dies einfach zu realisieren, indem das Projektorbild unverändert gelassen wird. Stereo-Kamerasysteme setzen dagegen eine Vorkenntnis der ungefähren Objektgeometrie im zu erfassenden Sichtfeld voraus, damit auch der Bildausschnitt der zweiten Kamera passend gewählt werden kann. Auch dies ist im Rahmen des hier vorgeschlagenen Verfahrens prinzipiell möglich, indem die in Kapitel 3 vorgeschlagene Repräsentation des teilvermessenen Objekts herangezogen wird. Im Verhältnis zum zusätzlichen Planungsaufwand ist der erzielbare Geschwindigkeitsvorteil bei Bildakquisition und Weiterverarbeitung (geringere Speicherbelastung) jedoch vermutlich bestenfalls marginal, so daß diese Idee nicht weiter verfolgt wurde.

4.3.3 Nebenbedingungen

Die zu planenden Aufnahmekonfigurationen unterliegen verschiedenen Nebenbedingungen, die sich aus dem Meßsystem selbst, seiner Interaktion mit der Umgebung, und aus der Weiterverarbeitung der Meßdaten ergeben. Sequeira [99] unterscheidet *operative Nebenbedingungen*, die sich aus dem Arbeitsbereich des Meß- und Positioniersystems ergeben, sowie *topologische Nebenbedingungen*, die sich aus der Geometrie der vermessenen Umgebung ergeben. Letztere sind auf den einfachen Sachverhalt zurückzuführen, daß der Sensor und das mit ihm verbundene Positioniersystem nur auf einer Seite der vermessenen Oberfläche (außerhalb des Objekts) positioniert werden können, nämlich dem einsehbaren Halbraum. Die eigentliche Kollisionsvermeidung zwischen Meßsystem und Objektoberfläche wird in der vorliegenden Arbeit ebenfalls den topologischen Nebenbedingungen zugerechnet (vgl. [99]).

Zusätzlich erfordert die Mehrbildintegration einen ausreichenden Überlappungsbereich, der zudem geometrisch so strukturiert sein muß, daß die Registrierung des neu akquirierten Bildes in den Bildverbund sicher möglich ist. Für typische Innenraumszenen und für viele Objekte ist diese zweite Forderung stets erfüllt; daher wurde diese

Nebenbedingung in dem im Rahmen der vorliegenden Arbeit vorgeschlagenen Verfahren nicht geprüft, sondern nur eine Mindestüberlappung gefordert. Im Falle von Objekten, die in großen Teilbereichen rotatorische oder translatorische Freiheitsgrade bei der Registrierung offenlassen, bietet sich eine vorherige Analyse mit den von Gelfand et al. [33] vorgeschlagenen Maßzahlen an, oder es muß auf photogrammetrische Paßmarken zurückgegriffen werden, wie sie zum Beispiel im Meßsystem von GOM mbH [34] Verwendung finden.

Planungstheoretisch sind die operativen Nebenbedingungen konstant, da sie ausschließlich von dem (unveränderten) Aufbau und der Kalibrierung des Meßsystems abhängen. Die topologischen Nebenbedingungen sowie die Nebenbedingung der Integrierbarkeit in den Mehrbildverbund sind dagegen in jedem Schritt k abhängig vom Informationsstand $i_k \in I$. Die in Abschnitt 3.3 vorgeschlagene Definition des eingesehenen Volumens und seiner Oberfläche liefert hierzu eine Abschätzung der Objektgeometrie, die sicher zur Berücksichtigung der topologischen Nebenbedingungen und zur Kollisionsvermeidung genutzt werden kann. Prinzipbedingt wird für noch nicht eingesehene Volumenbereiche angenommen, daß sie Teile des Objekts enthalten können, so daß falsch-positive Evaluationsergebnisse der operativen Nebenbedingungen nicht auftreten und Kollisionen sicher vermieden werden können.

4.4 Definition der Zielfunktion

Die in Abschnitt 4.3.1 formulierten Ziele sollen in diesem Abschnitt in eine formale Beschreibung des bei der Planung der jeweils nächsten Aufnahme zu optimierenden Kriteriums überführt werden. Während des Ablaufs der Vermessung wird dieses Kriterium nach jeder neuen Aufnahme und auf Basis der inkrementell ergänzten Objektbeschreibung genutzt, um die Aufnahmeparameter der nächsten Ansicht zu bestimmen. Die technische Umsetzung der Evaluierung des Kriteriums und der Optimierung des Parametersatzes bezüglich des Kriteriums ist Gegenstand von Abschnitt 4.5.

Der Grundgedanke des Kriteriums zur Bewertung möglicher Aufnahmekonfigurationen ist die einheitliche Behandlung der bereits vermessenen Oberflächenbereiche und der bis dahin nur durch unvermessene Grenzflächen des eingesehenen Volumens repräsentierten Teile des Objekts, die noch unvermessen sind. Für den bereits vermessenen Teil der Objektoberfläche kann die Auswirkung einer zusätzlichen Aufnahme zuverlässig prädiziert werden. Für die noch unvermessenen Bereiche der Objektoberfläche bieten die Grenzflächen des eingesehenen Volumens zwar keine quantitativ zuverlässigen Vorhersagemöglichkeiten, können aber aus mehreren Gründen als sinnvolle Ersatzrepräsentation mit zumindest qualitativ zielführendem Vorhersagepotential verwendet werden:

- Die Ersatzoberflächen verdecken stets unvermessene Bereiche der tatsächlichen Oberfläche. Wird also die Beobachtung einer Ersatzoberfläche geplant, so ist es wahrscheinlich, daß im Ergebnis ein unvermessener Teil der tatsächlichen Oberfläche eingesehen wird.
- Durch die geplante Beobachtung einer Ersatzoberfläche kann in jedem Fall eine Reduktion des bisher nicht eingesehenen Volumens erreicht werden.
- Die tatsächliche Lage der Oberfläche liegt stets *hinter* der Ersatzoberfläche, so daß die Grenzen des kalibrierten Meßbereichs zumindest in einer Richtung korrekt berücksichtigt werden können.

Wie schon aus der Formulierung dieser Argumente deutlich wird, hängt die Vorhersagegüte für die unvermessenen Bereiche von der korrekten Modellierung der Sichtbarkeit und der gegenseitigen Verdeckung ab. Obwohl in der vorliegenden Arbeit nur diskrete und opake Oberflächen mit entsprechend binären Sichtbarkeitsentscheidungen Verwendung finden, soll im Folgenden ein kurzer Exkurs zu den Möglichkeiten eingeschoben werden, die sich aus einer kontinuierlichen Sichtbarkeitsbewertung ergeben.

Exkurs kontinuierliche Sichtbarkeitsbewertung In der klassischen Verdeckungsrechnung opaker Objekte wird die Sichtbarkeit durch die Reihenfolge bestimmt, in der die Objektberandungen den Sichtstrahl schneiden; sichtbar ist der dem Betrachter nächstliegende Schnittpunkt, der die weiteren Oberflächen verdeckt. Auch im hier betrachteten Fall werden die Oberflächen als opak angenommen, aber ihre genaue Geometrie ist im Falle der Ersatzoberfläche für bisher nicht eingesehene Teile des Objekts unbekannt.

Die Unsicherheit, ob ein bestimmter Teil des Volumens opak oder aber leer und damit verdeckungsfrei ist, kann mit einem nicht-binären Sichtbarkeits-Koeffizienten modelliert werden. Die Umsetzung kann auf einfache Weise durch diskrete Oberflächenmodelle mit punkt- oder flächenelementweise spezifizierter Transparenz (α -Kanal) erfolgen. Aufwändigere Umsetzungen könnten eine räumliche Repräsentation der Opazitätswahrscheinlichkeit durch ein Volumenmodell berücksichtigen, das auch innerhalb des nicht eingesehenen Volumens unterschiedliche Transparenzen spezifiziert, beispielsweise proportional zum Abstand zum nächsten vermessenen Oberflächenpunkt.

Abbildung 4.1 veranschaulicht den Grundgedanken durch ein einfaches Beispiel. Nach einer ersten Aufnahme einer Säulenreihe mit dahinterliegender Wand ergeben sich hinter den Säulen uneingesehene Bereiche. Mit binärer Sichtbarkeitsentscheidung kann in diesem Fall kein substantieller Bewertungsunterschied zwischen den skizzierten Positionen 2a und 2b festgestellt werden. Die kontinuierliche Sichtbarkeitsbewertung ermöglicht dagegen die Modellierung des unbekannten Inhalts der verdeckten Volumina: mit einer gewissen Wahrscheinlichkeit befindet sich hinter der dritten

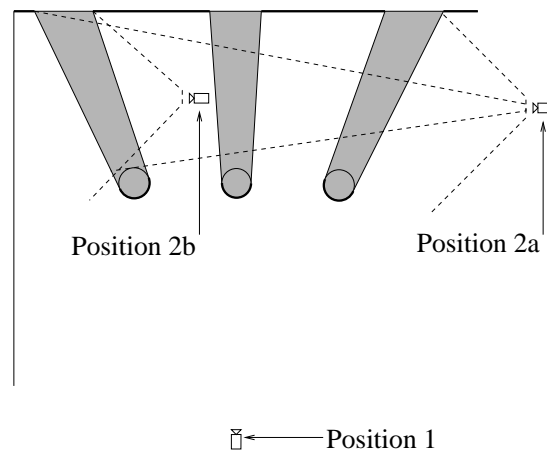


Abbildung 4.1: Die Aufnahmeplanung sollte Position 2a höher bewerten, da ein größerer Anteil unvermessener Oberflächen aufgelöst werden kann, wenn das initial verdeckte Volumen (grau) leer ist.

Säule eine vermessbare Wandfläche, die entsprechend in die Bewertung einbezogen wird, aber auch das gegenteilige Ereignis, welches die Vermessung einer durch die zweite (respektive durch die dritte) Säule verdeckten Wandfläche ermöglicht, kann mit entsprechender Gewichtung berücksichtigt werden.

Im Rahmen der vorliegenden Arbeit muß die kontinuierliche Sichtbarkeitsbewertung ein theoretischer Exkurs bleiben, da bis jetzt keine Möglichkeit einer schnellen Berechnung der entsprechenden Rendering-Schritte bei der Ergebnis-Prädiktion zu evaluierender Aufnahmeparameter besteht. Sowohl Volumen-Rendering als auch die reihenfolgeunabhängige Akkumulation von Transparenzen sind gegenwärtig im Vergleich zum Rendering opaker Oberflächen sehr viel rechenaufwändiger und verbieten sich daher aus Laufzeitgründen.

Hilfsfunktionen Zunächst sind die Definitionen einiger Hilfsfunktionen nützlich, um die Zielfunktion definieren zu können. Die Zielfunktion soll auf dem erwarteten Qualitätsgewinn gegenüber dem erreichten Zustand basieren und muß daher den vorliegenden, den gewünschten und den erwarteten Qualitätsstand jedes Oberflächenpunktes sowie seine Sichtbarkeit für einen gegebenen Satz Aufnahmeparameter berücksichtigen.

Die im vorangegangenen Exkurs thematisierte Sichtbarkeit wird für einen Punkt $x \in S$ auf der Objektoberfläche $S \in \mathbb{S}^2$ und einen gegebenen Satz $u \in U$ von Aufnahmeparametern (der insbesondere auch das optische Zentrum der Aufnahme enthält) durch die Funktion

$$w_{\text{sichtbar}} : \mathbb{R}^3 \times U \times \mathbb{S}^2 \rightarrow \{0, 1\} \quad (4.12)$$

beschrieben, mit $w_{\text{sichtbar}}(\mathbf{p}, u, \mathbf{S}_k) = 1$ wenn der Oberflächenpunkt \mathbf{p} mit dem Aufnahmeparametersatz u verdeckungsfrei vom Sensor innerhalb seines Blickfeldes eingesehen werden kann. Dabei wird der Informationsstand $i_k \in I_k$ des k -ten Iterationsschrittes zugrundegelegt, aus dem sich die Oberflächenrepräsentation \mathbf{S}_k ergibt. Ohne Beeinflussung der nachfolgenden Definitionen kann an dieser Stelle zunächst auch eine kontinuierliche Sichtbarkeitsbewertung mit $w_{\text{sichtbar}} : \mathbb{R}^3 \times U \times \mathbb{S}^2 \rightarrow \mathbb{R}$ angenommen werden; die in Abschnitt 4.5 beschriebene konkrete Umsetzung des Verfahrens geht aber von einer diskreten Sichtbarkeitsbewertung aus.

Zur Einbeziehung eines Qualitätskriteriums bei der Planung der nächsten Aufnahme muß zunächst der bereits erreichte Ist-Zustand der Vermessung durch die Datenqualität an jedem vermessenen Punkt der Oberfläche quantifiziert werden. Hierzu dient die Funktion

$$\tilde{\beta}_k : \mathbb{R}^3 \times I_k \rightarrow \mathbb{R} \quad (4.13)$$

welche mit Hilfe der k bisher erfolgten Aufnahmeaktionen $(u_1, \dots, u_k) \in \tilde{U}_k$ und der teilweise rekonstruierten Oberfläche $\mathbf{S}_k \in \mathbb{S}^2$ einem Punkt $\mathbf{p} \in \mathbf{S}_k \subset \mathbb{R}^3$ auf der bis dahin rekonstruierten Oberfläche ein Qualitätsmaß

$$\tilde{\beta}_k(\mathbf{p}, i_k) \text{ mit } i_k = (\mathbf{S}_k, u_1, \dots, u_k) \in I_k \quad (4.14)$$

zuordnet. Als Qualitätsmaß wird in der vorliegenden Arbeit die Abtastdichte der Meßpunkte auf der Oberfläche verwendet; für unvermessene Teile der Oberfläche ergibt das Qualitätsmaß daher den Wert 0. Zur konkreten Berechnung des Funktionswertes auf Basis der k registrierten Einzelaufnahmen sei auf Abschnitt 3.4.4 (Gleichungen 3.26, 3.33 und 3.34) verwiesen.

Im Unterschied zur Sichtbarkeitsfunktion w_{sichtbar} , die prinzipiell für beliebige Punkte im Raum evaluiert werden kann, kann das Qualitätsmaß nur für Punkte der vermessenen Oberfläche angegeben werden. Dies ermöglicht die zweidimensionale Speicherung als Attribut der Oberfläche, so daß im Kontext der Computergraphik die Repräsentation als (möglicherweise adaptiv aufgelöste) Texturkarte naheliegt.

Zusätzlich zum Ist-Zustand der erreichten Datenqualität wird auch der gewünschte Soll-Zustand durch das in jedem Abschnitt der Oberfläche zu erreichende Qualitätsmaß mit

$$\beta_{\text{max}} : \mathbb{R}^3 \rightarrow \mathbb{R} \quad (4.15)$$

angegeben. Der volumetrisch definierte Wertebereich ist an dieser Stelle notwendig, um auch initial unbekannten Objekten oder Umgebungen eine räumlich differenzierte Zielqualität zuweisen zu können. So kann die gewünschte Abtastdichte oftmals abhängig von der Höhe vorgegeben werden, um beispielsweise den Kopf einer Statue detaillierter zu erfassen, oder um die Decke eines Innenraums, die ohnehin außerhalb des relevanten Bereichs des eingesetzten mobilen Roboters liegt, nur grob zu vermessen.

Die durch die Parameter der evaluierten Aufnahme erwartete Qualität wird durch eine weitere Funktion

$$\beta_{k+1} : \mathbb{R}^3 \times U \times I_k \quad (4.16)$$

beschrieben, die analog zu Gleichung 4.13 einem Punkt $\mathbf{p} \in \mathbf{S}_k \subset \mathbb{R}^3$ auf der bis dahin rekonstruierten Oberfläche \mathbf{S}_k den Qualitätswert $\beta_{k+1}(\mathbf{p}, u, i_k)$ zuweist, der nach Umsetzung der Aktion $u \in U$ dem vorliegenden Informationsstand i_k zufolge mit Integration der evaluierten Aufnahme erreicht wird. Falls i_k bereits die vollständige Oberflächeninformation des zu rekonstruierenden Objekts enthält, so gilt der Zusammenhang

$$\beta_{k+1}(\mathbf{p}, u, i_k) \approx \tilde{\beta}_{k+1}(\mathbf{p}, i_{k+1}) \quad (4.17)$$

unter der Annahme, daß i_{k+1} den Informationsstand i_k ergänzt um das Ergebnis der Akquisitions-Aktion u beschreibt. Bei unvollständiger Oberflächeninformation ergeben sich erwartungsgemäß Abweichungen, insbesondere liegt der zur Evaluation von $\beta_{k+1}(\mathbf{p}, u, i_k)$ verwendete Punkt \mathbf{p} nach der $k + 1$ -ten Akquisition nicht notwendigerweise auf der rekonstruierten Oberfläche.

Auch für die Schätzung der durch den evaluierten Aufnahmeparametersatz erreichten Qualität wird in der vorliegenden Arbeit die Abtastdichte der Meßpunkte auf der Oberfläche verwendet, die über die Kamerageometrie (von der jeweiligen Meßrichtung abgedeckter Raumwinkel), die Entfernung zwischen Sensor und Oberfläche sowie die Oberflächennormale auf Basis der vorliegenden Oberflächenrepräsentation bestimmt werden kann. Für die konkrete Berechnung des Funktionswertes von β_{k+1} auf Basis der k registrierten Einzelaufnahmen und dem evaluierten Parametersatz u sei wiederum auf Abschnitt 3.4.4 verwiesen.

Abschließend ist noch eine Hilfsfunktion zur Bewertung der Kosten eines Satzes von Aufnahmeparametern zu definieren. Die Kostenfunktion soll im Wesentlichen verhindern, daß bei variabler Bildfeldgröße ein in zahlreiche sehr kleine Einzelaufnahmen fragmentierter Bildverbund entsteht, dessen Weiterverarbeitung entsprechend rechenaufwändig ist. Gleichzeitig soll bei variabler Bildauflösung das Datenvolumen als Kostenfaktor berücksichtigt werden. Mit der Funktion $c_p : U \rightarrow \mathbb{N}$, die dem Parametersatz einer Aufnahme die entstehende Pixelanzahl zuordnet, sowie den konstanten Kosten c_A einer zusätzlichen Bildakquisition und -integration ergibt sich

$$c(u) := c_p(u) + c_A \quad (4.18)$$

als Kostenfunktion. Weitere Kostenfaktoren wie beispielsweise die zurückzulegende Wegstrecke des Roboters von seiner aktuellen Position aus können an dieser Stelle leicht berücksichtigt werden; in der vorliegenden Arbeit wurde darauf verzichtet.

Zielfunktion Die zu optimierende Zielfunktion kann nun mit den im vorigen Abschnitt definierten Hilfsfunktionen formuliert werden. Dabei ist die jeweils zu erwartende Qualitätssteigerung, also die Differenz des Qualitätsmaßes vor und nach der Umsetzung der evaluierten Aktion, über der einsehbaren Oberfläche zu integrieren. Das Erreichen eines Qualitätsmaßes über die gewünschte Zielqualität hinaus soll die Bewertung der evaluierten Aktion nicht erhöhen.

Die Qualitätssteigerung an einem gegebenen Punkt $\mathbf{p} \in \mathbb{R}^3$ durch die zusätzliche Aufnahme mit Parametersatz $u \in U$ wird zur einfacheren Lesbarkeit zunächst zu einer Nutzenfunktion $B_k : \mathbb{R}^3 \times U \times I_k \rightarrow \mathbb{R}$ mit

$$B_k(\mathbf{p}, u, i_k) := \min\{\beta_{\max}(\mathbf{p}), \beta_{k+1}(\mathbf{p}, u, i_k)\} - \min\{\beta_{\max}(\mathbf{p}), \tilde{\beta}_k(\mathbf{p}, i_k)\} \quad (4.19)$$

zusammengefaßt. Da die erreichte Qualität durch Integration der zusätzlichen Aufnahme nicht abnehmen kann, gilt $B_k(x, u, i_k) > 0$ für alle $x \in S_k$ und alle $u \in U$. Damit ergibt sich unter Berücksichtigung von Kosten und Nutzen des evaluierten Parametersatzes $u \in U$ eine Zielfunktion $G : U \times I_k \rightarrow \mathbb{R}$ mit

$$G(u, i_k) := \frac{c_p(u) + c_A}{\iint_{\mathbf{p} \in S_k} w_{\text{sichtbar}}(\mathbf{p}, u, S_k) \cdot B_k(\mathbf{p}, u, i_k) \, dS} \quad (4.20)$$

und $i_k = (S_k, u_1, \dots, u_k)$, welche die oben gestellten Anforderungen erfüllt. Den im Nenner berücksichtigten Kosten wird im Zähler die prädierte Qualitätssteigerung im Integral über dem sichtbaren Teil der Objektoberfläche gegenübergestellt.

Auf diese Weise kann die gesuchte Aufnahmesequenz u_1, \dots durch einen *greedy*-Ansatz iterativ bestimmt werden. In der Praxis lassen sich die relevanten Basisdaten zur Berechnung der benötigten Funktionswerte für jeden Punkt der Oberfläche S auf

- Sichtbarkeit
- räumliche Lage
- Oberflächennormale
- und ein Qualitätsmaß für vorangegangene Aufnahmeschritte

zurückführen. Im folgenden Abschnitt wird ein Verfahren vorgeschlagen, mit dem das Integral in Gleichung 4.20 auf effiziente Weise mit der in der Praxis benötigten Genauigkeit berechnet werden kann.

4.5 Effiziente Berechnung der Zielfunktion

Die Berechnung der im vorangegangenen Abschnitt in Gleichung 4.20 vorgeschlagene Zielfunktion erscheint in der praktischen Umsetzung auf den ersten Blick schwierig. Mit der numerischen Standardmethode der Summierung von Stichproben müssen

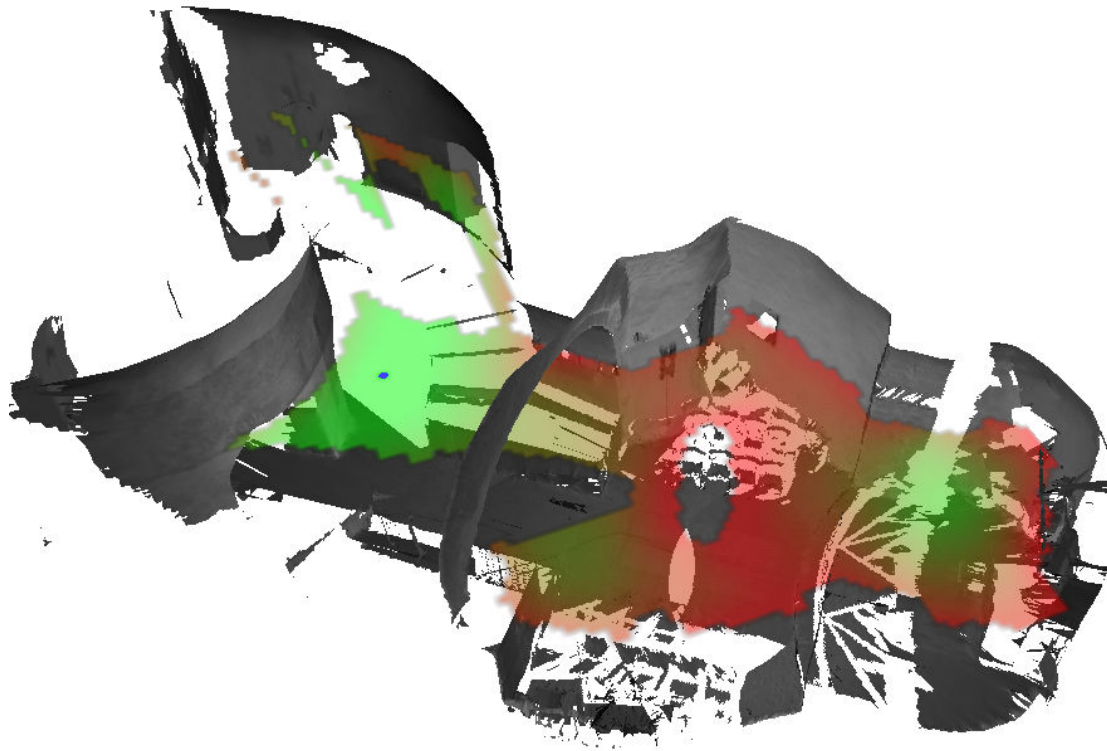


Abbildung 4.2: Farbkodierte Darstellung der Zielfunktionswerte für einen omnidirektional arbeitenden Lichtlaufzeitscanner, für den nur die Aufnahmeposition innerhalb einer Ebene (hier transparent dargestellt) variiert wurde. Rot bezeichnet hohe, grün niedrige Funktionswerte, die optimale nächste Aufnahmeposition ist durch einen blauen Punkt gekennzeichnet. Die verwendete Datensituation entstammt der Vermessung einer dänischen Rundkirche, die auch in Abbildung 4.5 zu sehen ist.

auf der integrierten Oberfläche S zunächst in geeigneter Dichte Stichprobenpositionen definiert werden. Für jede dieser Stichprobenpositionen muß zunächst unter Berücksichtigung aller übrigen Teile der Oberflächenrepräsentation die Verdeckungsrechnung bezüglich der evaluierten Aufnahmeposition durchgeführt werden, bevor die übrigen Eigenschaften der Stichprobenposition in die Integralberechnung einfließen können. Der in der vorliegenden Arbeit vorgeschlagene Lösungsansatz zu dieser Problematik basiert auf der Analogiebildung zwischen der Tiefenbildakquisition mit einem 3D-Meßsystem und dem computergraphischen *rendering*, also der Überführung einer bereits vorliegenden 3D-Szenenbeschreibung in ein Bild.

4.5.1 Grundprinzip

Die in Gleichung 4.20 vorgeschlagene Zielfunktion abstrahiert von den tatsächlich eingesetzten 3D-Meßsystemen. Für ihre praktische Berechnung ist es sinnvoll, die zugrundeliegende Idee einer rechnerischen Ergebnis-Prädiktion der jeweils evaluierten Aufnahmekonfiguration als zentralen Ansatz zu wählen. Die im Bildraster der jeweiligen Aufnahme organisierten Meßpunkte liefern bereits eine sinnvolle Verteilung der Stichprobenpositionen zur numerischen Integralberechnung. Daher ist es naheliegend, zur evaluierten Aufnahmekonfiguration zunächst ein Bild zu erzeugen, das die sichtbaren Punkte der Oberfläche S enthält.

Die Akquisition von Tiefenbildern einer realen Szene mit einem 3D-Meßsystem entscheidet die Sichtbarkeit der im Bild zu repräsentierenden Oberfläche inhärent auf eindeutige Weise. Transparente oder semitransparente Gegenstände, die diese Eigenschaft konterkarieren würden, müssen durch entsprechende Vorbereitung (zum Beispiel Einsprühen) opak gemacht oder aus der Szene entfernt werden. Im Ergebnis wird das Blickfeld des 3D-Meßsystems so auf einem zweidimensional geordneten Bündel von Sichtstrahlen optisch abgetastet, daß der nächstliegende Schnittpunkt des jeweiligen Sichtstrahls mit der Oberfläche als Meßpunkt erfaßt wird.

Der duale Prozess in der Computergraphik ist das *rendering*, bei dem ebenfalls zu einem zweidimensional organisierten Bündel von Sichtstrahlen die nächstliegenden Schnittpunkte mit der bereits vorliegenden 3D-Oberfläche gefunden werden müssen. Die wichtigsten bildpunktweise arbeitenden Verfahren sind hier *ray-casting* (Strahlverfolgung) und der *z-buffer*-Algorithmus. Letzterer ist seit langem in geschwindigkeitskritischen Anwendungen als Standardverfahren etabliert; effiziente hardware-basierte Implementierungen sind in fast allen Graphikkarten anzutreffen einschließlich kostengünstiger Modelle für den Massenmarkt.

Im Allgemeinen wird es allerdings nicht effizient möglich sein, die Bildgeometrie des 3D-Meßsystems mit allen bei der Umrechnung der Meßwerte in 3D-Koordinaten zu korrigierenden Eigenschaften (zum Beispiel Verzerrung) computergraphisch nachzubilden. Insbesondere die Aufnahmen von 360°-Systemen mit sphärischer Bildgeometrie lassen sich mit den auf perspektivische und telezentrische Bildgeometrien spezialisierten marktüblichen Graphiksubsystemen nur mit hohem Aufwand modellieren. Dies ist für das vorgeschlagene Verfahren allerdings auch nicht notwendig: die graphische Darstellung muß lediglich ein Bild liefern, welches *homöomorph* zum Bild des 3D-Meßsystems ist. Formal ist diese Forderung äquivalent zur Geltung der folgenden Eigenschaften der Tiefenbilder $r(u, v)$ des 3D-Meßsystems und $r'(u', v')$ der computergraphischen Bildgeometrie:

- Die Bilder $r(u, v)$ und $r'(u', v')$ besitzen umkehrbare Projektionsfunktionen $h : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ und $h' : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, die einem Bildpunkt $(u, v, r(u, v))^T$ eine Position im Raum $h^{-1}((u, v, r(u, v))^T)$ zuordnen (h' entsprechend).

- Es existiert eine Abbildungsfunktion $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ mit $f = h^{-1} \circ h'$.
- f ist bijektiv und stetig.
- Es existiert eine Umkehrfunktion f^{-1} , die ebenfalls stetig ist.

Diese Bedingungen sind für Bildgeometrien, die sich auf den gleichen Blickpunkt beziehen, erfüllt. Daher kann eine computergraphisch generierte Ansicht verwendet werden, um die für die numerische Integralberechnung benötigten Stichprobenpositionen auf der Oberfläche S zu gewinnen. Dieses Grundprinzip wird in Abbildung 4.3 veranschaulicht.

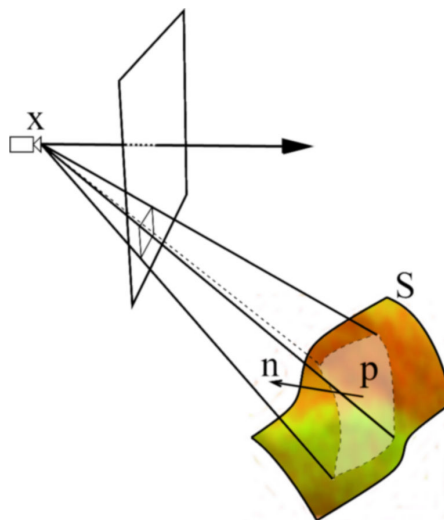


Abbildung 4.3: Geometrie des computergraphischen Prädiktionsverfahrens. Das Bildraaster (exemplarisch durch einen Pixel in der Bildebene dargestellt) liefert die Summenelemente für die numerische Integration. Mit dem bekannten Raumwinkel, der Entfernung und dem Winkel zwischen Oberflächennormale und Blickrichtung wird die Abtastdichte der Messung präzisiert. Die Farbe kodiert die bereits erreichte Qualität.

4.5.2 Abbildung der Zielfunktion auf graphische Basisfunktionen

4.5.2.1 Erwartete Abtastdichte

In Abschnitt 3.4.4.1 wurde in Gleichung 3.26 gezeigt, wie die Abtastdichte durch lokale Operatoren im Bildraaster mit der Projektionsfunktion h berechnet werden kann. Im Zusammenspiel mit computergraphisch generierten Ansichten stellt sich die Frage, ob die Anwendung lokaler Operatoren zur Bestimmung der Oberflächennormale

nicht vermieden werden kann, indem sie durch Auswertung graphischer Standardrechenschritte ersetzt wird, wie sie zum Beispiel bei der Berechnung der Schattierung auftreten.

Der relevante Rechenschritt beim Verfahren der Schattierung ist die Berechnung des reflektierten Lichtes von einer Oberfläche, die als diffus reflektierend angenommen wird, nach dem Lambert'schen Kosinus-Gesetz. Dabei wird der Kosinus des Winkels zwischen der Einfallsrichtung des Lichts \vec{c} (Vektor vom betrachteten Oberflächenpunkt zur Lichtquelle) und der Oberflächennormale \vec{n} bestimmt. Ein ähnlicher Rechenschritt wird auch bei der Berechnung des Raumwinkels⁵ durchgeführt, den ein gegebener Oberflächenabschnitt $s \subset S$ abdeckt, wenn er vom Punkt x aus betrachtet wird:

$$A = \iint_{p \in s \subset S} \frac{\cos((\mathbf{p} - \mathbf{x}) \cdot \mathbf{n}|_p)}{(\mathbf{p} - \mathbf{x})^2} dS \quad (4.21)$$

Unter der Annahme kleiner und näherungsweise linearer Oberflächenabschnitte kann mit dieser Beziehung auch umgekehrt der Flächeninhalt eines Oberflächenabschnitts abgeschätzt werden, wenn Raumwinkel, Entfernung und der gemittelte Kosinus-Ausdruck bereits bekannt sind. Dies ist gerade die Datensituation, die in einem Bildraster vorliegt, in welchem jedem Pixel bereits ein von ihm abgedeckter Raumwinkel A sowie ein Entfernungswert (z -buffer-Eintrag) und ein Grauwert aus der schattierten Darstellung der Oberfläche zugeordnet ist. Die Schattierung ist dabei das Ergebnis eines Beleuchtungsmodells, das eine einzige Punktlichtquelle in der Beobachtersposition x sowie eine lambert-reflektierende Oberfläche beschreibt. Die Abtastdichte d des Bildrasters auf der Oberfläche ergibt sich in diesem Fall durch

$$d = \frac{4\pi \cos((\mathbf{p} - \mathbf{x}) \cdot \mathbf{n})}{A \cdot (\mathbf{p} - \mathbf{x})^2} \quad (4.22)$$

oder, ausgedrückt mit der Schattierung s , dem z -buffer-Eintrag r und dem Raumwinkel A zur jeweiligen Pixelposition (u, v) , durch

$$d = \frac{4\pi \cdot s(u, v)}{A(u, v) \cdot r(u, v)^2} \quad (4.23)$$

Zweckmäßigerweise wird man den Wert $4\pi \cdot A(u, v)^{-1}$ für jeden Pixel (u, v) einmalig vorberechnen, während $s(u, v)$ und $r(u, v)$ (bis auf Offset und Skalierung) für jede evaluierte Akquisitions-Aktion vom Graphiksubsystem berechnet werden. Die Dekodierung der Schattierungswerte wird im folgenden Unterabschnitt 4.5.2.2 beschrieben.

⁵Das Maß des Raumwinkels ist (zitiert nach Bronstein [12]) „die Fläche, die von diesem Raumwinkel aus der Einheitskugel um den Scheitel als Mittelpunkt herausgeschnitten wird“, liegt also zwischen 0 und 4π . Der Scheitel ist im vorliegenden Fall der Beobachtungspunkt \vec{x} .

Relevant für die Berechnung der Zielfunktion ist an dieser Stelle allerdings nicht die oben beschriebene Berechnung der Abtastdichte durch das Bildraster der computergraphischen Darstellung, sondern die der Abtastdichte des realen Sensors. Wenn die computergraphische Ansicht, wie in Abschnitt 4.5.1 gefordert, homöomorph zur Bildfeldgeometrie des 3D-Meßsystems ist, so kann für jeden Pixel der computergraphischen Ansicht auch die Anzahl der Meßpunkte bestimmt werden, die innerhalb des vom computergraphischen Pixel abgedeckten Raumwinkels durch das 3D-Meßsystem akquiriert werden. Die Details des rechnerischen Zusammenhangs zwischen der Auflösung des Sensors und der computergraphischen Ansicht werden im Anhang in Abschnitt A.1 dargestellt.

4.5.2.2 Erreichte Abtastdichte

Durch die im vorigen Unterabschnitt vorgeschlagene computergraphische Berechnung der erwarteten Abtastdichte unter Verwendung der schattierten Darstellung bleiben von den drei verfügbaren Freiheitsgraden der RGB-Darstellung zwei übrig, die zur Codierung von Qualitätsmaßen genutzt werden können. Im Rahmen der vorliegenden Arbeit wurde das Verhältnis von Grün zur Summe der Farbkanäle genutzt, um die bereits erreichte Abtastdichte zu codieren. Für den Benutzer ergibt sich damit eine intuitiv leicht verständliche Darstellung, die vorgabengerecht abgetastete Oberflächenbereiche grün, unzureichend abgetastete Bereiche dagegen rot anzeigt.

Bei der Definition der Graphikprimitive wird die erreichte Qualität q im Verhältnis zur maximal gewünschten Qualität q_{max} durch das RGB-Farbtupel (r, g, b) mit

$$\begin{aligned} r &= (1 - g - b) \\ g &= q/q_{max} \\ b &= 0 \end{aligned} \tag{4.24}$$

codiert. Die Werte für die erreichte Qualität q und für die Schattierung s , die in Gleichung 4.23 Verwendung findet, lassen sich aus dem Farbtupel (r, g, b) durch

$$q = \frac{g}{r + g + b} \cdot q_{max} \tag{4.25}$$

und

$$s = r + g + b \tag{4.26}$$

rekonstruieren.

4.5.2.3 Numerische Integration

Aus Gleichung 4.23 wird zunächst die abgedeckte Fläche $F_{graphic}$ der Objektoberfläche mit

$$F_{graphic}(u, v) = \frac{A_{graphic}(u, v) \cdot r(u, v)^2}{4\pi \cdot s(u, v)} \quad (4.27)$$

abgeleitet, die einem Pixel der computergraphischen Darstellung zugeordnet werden kann. Mit den in Gleichungen 4.23 und 4.25 definierten prädizierten Qualität $d(u, v)$ und erreichten Qualität $q(u, v)$ zum Pixel (u, v) kann nun die Summe über der im Bild sichtbaren Oberfläche zur Approximation des Integrals aus Gleichung 4.20 durch

$$\begin{aligned} \iint_{\mathbf{p} \in \mathbf{S}_k} w_{sichtbar}(\mathbf{p}, u, \mathbf{S}_k) \cdot B_k(\mathbf{p}, u, i_k) \, d\mathbf{S} \\ \approx \sum_{u, v} F_{graphic}(u, v) \cdot (\min\{\beta_{max}, d(u, v)\} - \min\{\beta_{max}, q(u, v)\}) \end{aligned} \quad (4.28)$$

gebildet werden.

4.5.2.4 Omnidirektionale Bildrepräsentation

Verglichen mit der computergraphischen Berechnung der sichtbaren Oberfläche ist die weitere Verarbeitung der generierten Daten nach der in Gleichung 4.23 pixelweise definierten Vorgehensweise sehr viel weniger rechenaufwändig. Daher stellt sich die Frage, wie die erneute Berechnung der sichtbaren Oberfläche vermieden werden kann, wenn zum gleichen Beobachtungspunkt verschiedene Blickrichtungen, Blickfelder und Bildauflösungen evaluiert werden sollen.

Hierzu ist theoretisch wie methodisch ein Blick auf die Verfahren des sogenannten *image based renderings* sinnvoll. Einer der zentralen Begriffe des *image based renderings* ist die von Adelson und Bergen [1, zitiert nach [67]] definierte *plenoptische Funktion*. Diese ordnet einem Blickpunkt \mathbf{x} , einem horizontalen und vertikalen Blickwinkel (ϑ, φ) und einem betrachteten Wellenlängenbereich λ einen Intensitätswert i mit

$$i = P(\vartheta, \varphi, \lambda, \mathbf{x}) \quad (4.29)$$

zu (die in der Originalarbeit enthaltene Parametrisierung nach der Zeit wird hier weggelassen, da nur statische Szenen betrachtet werden). McMillan [67] definiert eine Stichprobe der plenoptischen Funktion als das Abbild der Umgebung zu einem gegebenen Blickpunkt. Das primäre Ziel des *image based renderings* besteht nun darin, aus mehreren Stichproben durch geeignete Interpolationsmethoden neue Ansichten der repräsentierten Szene zu gewinnen.

Ebenso wie beim *image based rendering* muss auch im Rahmen der vorliegenden Arbeit die räumliche Anordnung der Stichproben der plenoptischen Funktion so gewählt werden, daß eine hinreichend genaue Beschreibung der zu evaluierenden Ansichten entsteht. Anstelle synthetisierter Ansichten genügt für die Aufnahmeplanung aber die Bestimmung des in Gleichung 4.20 definierten Zielfunktionswertes, der als Skalar auch eine wesentlich einfachere Interpolation zwischen den Stichproben (Stützstellen) ermöglicht.

Um die oben genannten Vorteile – Vermeidung der erneuten Verdeckungsrechnung bei Evaluation mehrerer Ansichten zu gleichen Blickpunkt – mit dem Ansatz der plenoptischen Funktion nutzen zu können, muß diese in einem geeigneten Raster abgetastet und als Bild gespeichert werden. Dabei ist im Zusammenspiel mit der Erzeugung der repräsentierten Daten durch die Graphikhardware eine Bildgeometrie sinnvoll, die mit geringem Aufwand berechnet werden kann. Zylindrische und sphärische Bildgeometrien scheiden damit aus.

Gut geeignet ist dagegen die von Greene [35] erstmalig vorgeschlagene Repräsentation durch die sechs Flächen eines Würfels. Analog zur von Cohen und Greenberg im Kontext der *radiosity*-Rechnung [17] eingeführten Bezeichnung “hemi cube” für die Repräsentation des Halbraums über einem Oberflächenpunkt wird in der vorliegenden Arbeit der Begriff “view cube” für die computergraphisch erzeugte Repräsentation der von einem Raumpunkt aus einsehbaren Umgebung vorgeschlagen (Abbildungen 4.4 und 4.5).

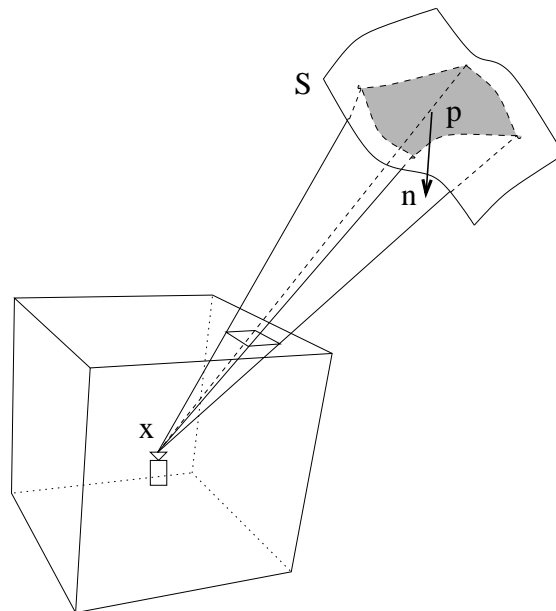


Abbildung 4.4: Der view-cube mit der von einem Pixel abgedeckten Objektoberfläche.

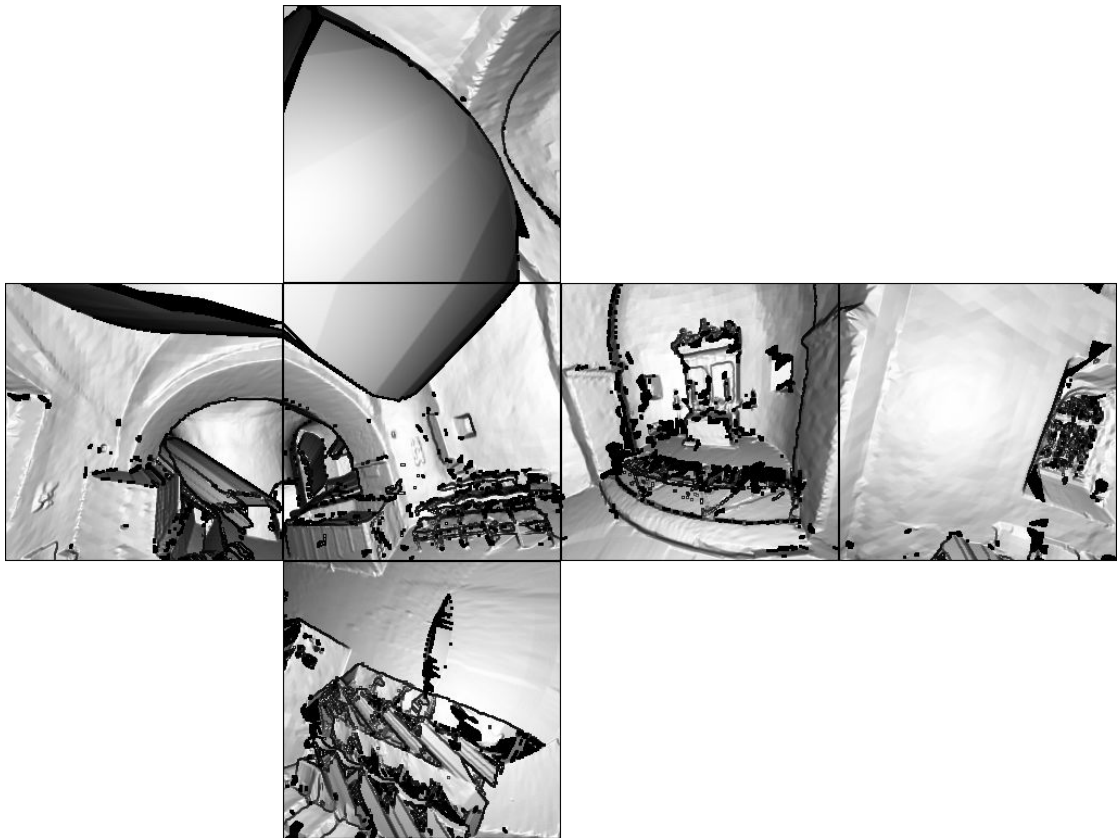


Abbildung 4.5: *Schattierter Inhalt des view-cubes für die Planung der dritten Aufnahme einer dänischen Rundkirche.*

Aus implementierungstechnischer Sicht bildet der “view cube” in der Berechnung der Zielfunktion eine Tabelle mit Zwischenergebnissen, die computergraphisch erzeugt wird. Hierzu können als computergraphische Primitive sowohl Einzelpunkte im Punkterendering eingesetzt werden als auch Dreiecksnetze unter Anwendung der Gouraud-Schattierung, sofern die Dreieckskanten mit Unstetigkeiten in der Abtastdichte hinreichend genau koinzidieren.

Die Ergebnisse der vorliegenden Arbeit zur geplanten internen 3D-Rekonstruktion wurden mit Dreiecksnetzen erzeugt, die einer Zwischenstufe des *mesh-zippering*-Verfahrens entnommen werden. Dabei wird zu jeder Einzelaufnahme ein Teilnetz erzeugt und von Redundanzen mit anderen Teilnetzen bereinigt. Auf diese Weise können zwei wichtige Forderungen der computergraphischen Berechnung der Zielfunktion erfüllt werden:

- Suboptimale Teilmessungen sind nicht repräsentiert, so daß als Kombinations-

verfahren in redundant vermessenen Bereichen das Maximum der erreichten Einzelabtastdichten angenommen wird.

- Unstetigkeiten der Abtastdichte auf der Oberfläche sind bereits durch die Entfernung suboptimal vermessener Bereiche aus jedem Teilnetz ausgeschlossen.

4.5.2.5 Stereoskopische Sichtbarkeit

Die Vermessbarkeit eines Objekts mit einem triangulationsbasierten 3D-Meßsystem unterscheidet sich grundsätzlich von einem unifokalen Meßsystem, da die Oberfläche nur dann erfaßt wird, wenn sie von mehreren, meist zwei, verschiedenen Beobachtungspunkten aus eingesehen werden kann. In der akquirierten Einzelaufnahme ergeben sich die für triangulationsbasierte Systeme typischen Abschattungen, die vom ungeübten Benutzer nur schwer vorherzusehen sind.

In der Prädiktion der Meßsituation ist das ein wichtiger Effekt, der in der computergraphischen Umsetzung mit etablierten Verfahren der Schatten-Rechnung [44, 115, 23] erzeugt werden kann. Dabei kann der primäre Beobachtungspunkt auch in der computergraphischen Darstellung als Beobachtungspunkt verwendet werden, während der optisch aktive oder auch passive sekundäre Beobachtungspunkt (Projektor oder zweite Kamera) wie eine Punktlichtquelle in der Schatten-Rechnung berücksichtigt wird. Alle sonstigen aus dem Bereich der *radiosity* bekannten Effekte wie weitere Lichtquellen oder indirekte Beleuchtung entsprechen in der hier betrachteten Meßaufgabe einer Störung der Messung, die unabhängig von der Planung soweit minimiert werden sollte, daß sie vernachlässigbar wird. Das einzusetzende Verfahren ist daher die Darstellung von Schlagschatten (engl. *hard shadows*), die glücklicherweise weniger aufwändig zu berechnen sind als die in natürlichen Szenen typischerweise auftretenden Halbschatten⁶.

Die Berechnung der Stereo-Ansicht erfolgt in einem zweistufigen *rendering*-Verfahren, in dessen erstem Schritt die Szene aus der Sicht des sekundären Beobachtungspunktes gerendert wird. Die dabei berechnete Tiefenkarte (*z buffer*) wird als Schattenkarte bezeichnet. Die Darstellung der Szene aus der Sicht des primären Beobachtungspunktes erfolgt für jeden Bildpunkt mit einem zusätzlichen Test auf Abschattung:

- Sichtbarkeitsprüfung durch *z-buffer*-Test.
- Abbildung des Bildpunktes in die Schatten-Karte, dort erneuter Sichtbarkeitsprüfung durch *z-buffer*-Test.

⁶Durch Verunschärfung beziehungsweise durch die Apertur der Optik zum sekundären Beobachtungspunkt können prinzipiell auch in triangulationsbasierten Meßsystemen Halbschatten auftreten. Dabei wirkt aber die abschattende Geometrie wie eine zusätzliche optische Blende, so daß die im Halbschatten liegenden Bereiche in aller Regel nicht maßhaltig erfaßt werden können.

- Schattierte Darstellung, wenn der Objektpunkt von beiden Blickpunkten eingesehen werden kann, sonst Darstellung in Hintergrundfarbe.

Die hier skizzierte Berechnung kann mit geringem Aufwand zum Beispiel mit standardisierten Erweiterungen⁷ der OpenGL [98, 58] unter Ausnutzung hardwarenaher Implementierungen umgesetzt werden [30]. Es sei noch angemerkt, daß die bekannten Aliasing-Probleme der Schattenrechnung in der geometrischen Konfiguration triangulationsbasierter Meßsysteme typischerweise nicht auftreten können, da beide Beobachtungspunkte ungefähr den gleichen Abstand und ähnliche Winkel zum Objekt haben.

Anders als im Falle der Punktlichtquelle ergeben sich im Falle eines triangulationsbasierten Meßsystems allerdings neben der einfachen Abschattung weitere Fälle, in denen die Oberfläche nicht erfaßt werden kann. Diese Einschränkungen werden durch die seitlichen Grenzen des auswertbaren Meßvolumens sowie durch den kalibrierten Tiefenmeßbereich zu den beiden Blickpunkten verursacht.

Während sich die Berücksichtigung der seitlichen Grenzen des Meßvolumens für den primären Beobachtungspunkt trivial durch das gerenderte Bildfeld ergibt, erfordern die übrigen Grenzen des Meßvolumens eine explizite Berücksichtigung. In der vorliegenden Arbeit werden hierzu die Grenzflächen des Meßvolumens durch ein Dreiecksnetz modelliert, das in Abbildung 4.6 graphisch dargestellt ist. Dies ist meist mit

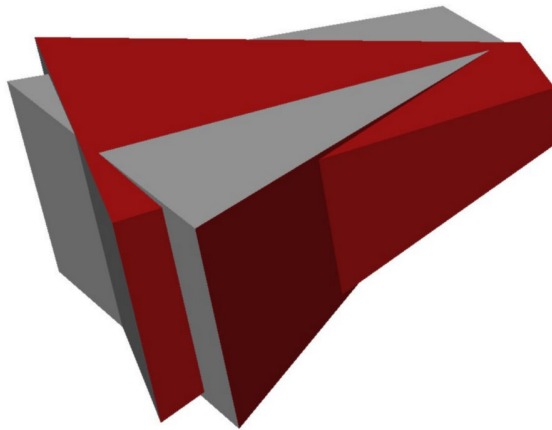


Abbildung 4.6: 3D-Bildfeldgeometrie eines realen Triangulationsmeßsystems mit jeweils 22.6° horizontalem Öffnungswinkel und 18.8° Triangulationswinkel. Die Flächen in A_{out} sind grau, die in A_{in} sind rot dargestellt. Der primäre Beobachtungspunkt, auf den sich die Flächenklassifizierung bezieht, liegt in der dargestellten Ansicht vorne.

⁷Die relevanten Erweiterungen `GL_ARB_shadow` beziehungsweise `GL_SGIX_SHADOW` zur einfachen Darstellung von Schatten mit OpenGL sind bereits seit einigen Produktgenerationen als hardwarebeschleunigte Implementierungen der einschlägigen Hersteller verfügbar: Silicon Graphics ab Infinite Reality 2, NVIDIA ab GeForce3, ATI ab Radeon9500 [44].

wenig Aufwand möglich, da es sich bei der typischen Blickfeldgeometrie von Projektoren und Kameras um Pyramidenstümpfe handelt. Krümmungen der Randfläche aufgrund von Verzeichnungen können an dieser Stelle auf einfache Weise berücksichtigt werden; bei den typischerweise anzutreffenden Kameralinsen kann dieses Detail aber vernachlässigt werden. Eine Fusion der Blickfeldgeometrien von primärem und sekundärem Beobachtungspunkt ist bei dem nachfolgend vorgeschlagenen Verfahren nicht erforderlich.

Für die konkrete Prädiktion des vermeßbaren Bereichs der Oberfläche wird das die Grenzfläche des Meßvolumens beschreibende orientierte Polyeder (Dreiecksnetz) zunächst in zwei disjunkte Teilmengen zerlegt. Der Teil der Grenzfläche, der vom primären Beobachtungspunkt aus von außen sichtbar ist wird als A_{in} , der übrige Teil dagegen als A_{out} bezeichnet. Mit diesen Zusatzflächen geschieht die Darstellung in folgenden reihenfolgeabhängigen Teilschritten:

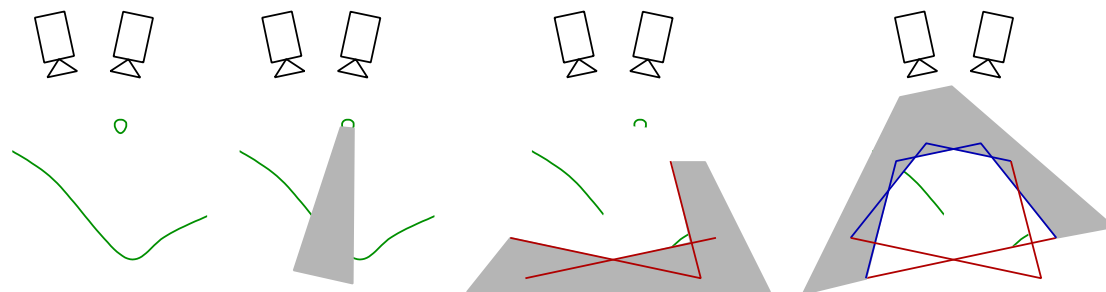


Abbildung 4.7: Sichtbarkeits-Bestimmung für ein Stereo-Meßsystem bei der Vermessung einer Objektoberfläche (grün). Die Flächen in A_{out} sind rot, die in A_{in} sind blau dargestellt. Der laterale Beschnitt am Bildrand (clipping, nicht z-buffering) der primären (linken) Kamera wird hier graphisch nicht unterschieden.

1. Berechnung der Schatten-Karte.
2. Darstellung der Objektrepräsentation mit Test auf Abschattung.
3. Darstellung von A_{out} ohne Test auf Abschattung in der Hintergrundfarbe. Durch den *z-buffer*-Tests wird der Teil der Objektrepräsentation, der aus Sicht der Kamera hinter A_{out} liegt, mit der Hintergrundfarbe überzeichnet.
4. Darstellung von A_{in} mit invertiertem *z-buffer*-Test ohne Test auf Abschattung in der Hintergrundfarbe. Durch den invertierten *z*-Vergleich wird der Teil der Objektrepräsentation, der aus Sicht der Kamera vor A_{in} liegt, mit der Hintergrundfarbe überzeichnet.

Im Ergebnis werden gerade die Oberflächen dargestellt, die vom 3D-Meßsystem zuverlässig (durch Beschränkung auf das kalibrierte Meßvolumen) vermessen werden können. Die übrigen Teile des Objekts werden dennoch bei der Sichtbarkeitsprüfung

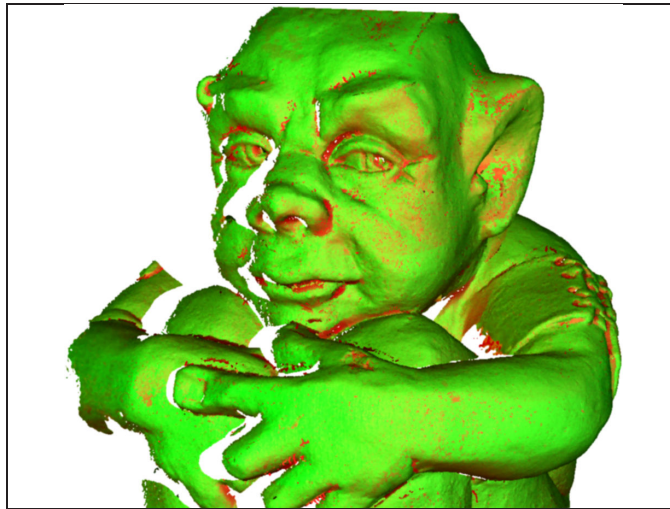


Abbildung 4.8: Sichtbarkeits-Bestimmung für ein Stereo-Meßsystem durch Methoden der Schattenrechnung. Der zweite Beobachtungspunkt liegt rechts vom Beobachtungspunkt des dargestellten Bildes und verursacht unvermessbare Bereiche im Schatten der Nase. Am oberen Bildrand ist die Grenze des Sichtfeldes des zweiten Beobachtungspunktes zu erkennen.

und bei der Schattenrechnung berücksichtigt. Das Verfahren ist in Abbildung 4.7, ein typisches Ergebnis in Abbildung 4.8 graphisch dargestellt. Durch die exakte Vorhersage des tatsächlich vermessbaren Oberflächenbereichs ist das Verfahren auch bei der Gestaltung einer graphischen Benutzerschnittstelle zur manuellen Robotersteuerung hilfreich.

Kombination mit der omnidirektionalen Bildrepräsentation Angesichts des vergleichsweise aufwändigen zweistufigen *rendering*-Verfahrens stellt sich die Frage, ob dieses durch Anwendung der im vorigen Unterabschnitt vorgeschlagenen omnidirektionalen Bildrepräsentation beschleunigt werden kann. Als Ansatzpunkte können zunächst zwei Fälle unterschieden werden:

Rotation um die Basislinie: In diesem Fall bleiben sowohl primärer als auch sekundärer Beobachtungspunkt unverändert. Die eingesehene Umgebung kann daher analog zum *view cube* auf vier Bildflächen, die topologisch einem Zylinder entsprechend angeordnet sind, zwischengespeichert werden.

Rotation um einen der Beobachtungspunkte: In diesem Fall kann die Schattenkarte omnidirektional vorberechnet und zur Darstellung mehrerer Ansichten aus der Sicht des primären Beobachtungspunktes verwendet werden.

Alternativ kann das von Zhang vorgeschlagene Verfahren des *forward shadow*

mappings [118, zitiert nach [119]] verwendet werden, bei dem die Schatten-Karte als $2\frac{1}{2}$ D-Textur auf das Bild des primären Beobachtungspunktes aufprojiziert wird. In diesem Fall kann der *view cube* zur omnidirektionalen Zwischenspeicherung der einsehbaren Umgebung des primären Beobachtungspunktes verwendet werden.

In allen Fällen müssen die Grenzflächen des Meßvolumens A_{in} und A_{out} zu jeder evaluierten Akquisitionsaktion nachträglich berücksichtigt werden.

In der Praxis der 3D-Rekonstruktion von Objekten stellt sich allerdings heraus, daß ein sehr großer Teil der Ansichten konvexe Oberflächenbereiche zeigt, so daß einer Variation der Rotation um die Basislinie im Vergleich zur Rotation um den Objektmittelpunkt oder um den primären Beobachtungspunkt nur eine untergeordnete praktische Bedeutung zukommt. Auch für den relevantesten Fall der Rotation um die Blickachse der primären Kamera kann nur eine geringe Beschleunigung der Aufnahmeplanung (weniger als 50%) erreicht werden. Daher wurde die Kombination der omnidirektionalen Bildrepräsentation mit der stereoskopischen Sichtbarkeitsprüfung im Rahmen der vorliegenden Arbeit nicht praktisch umgesetzt.

4.6 Globale Optimierung der Aufnahmeparameter

4.6.1 Anforderungen

Nachdem die Zielfunktion der iterativen Aufnahmeplanung in den vorangegangenen Abschnitten definiert und ein Verfahren zu ihrer effizienten Berechnung vorgeschlagen wurde, soll nun die Algorithmik zur Bestimmung der jeweils nächsten Akquisitionsaktion beschrieben werden. Dabei handelt es sich um ein kontinuierliches globales Optimierungsproblem mit Nebenbedingungen. Durch die Berechnung der Zielfunktion mit der eingeschränkten Bilddiskretisierung wird die Lösung dieses Problems zusätzlich durch Rauschen erschwert. An das Optimierungsverfahren werden folgende Anforderungen gestellt:

- Zwischenwerte, die als omnidirektionale Bildrepräsentation im *view cube* gespeichert werden, sollen optimal ausgenutzt werden. Daher muß das Optimierungsverfahren eine geeignet geordnete Änderung der einzelnen Parameter unterstützen, so daß insbesondere der Beobachtungspunkt für möglichst viele aufeinanderfolgende Evaluationen unverändert bleibt.
- Das Verfahren muß robust gegenüber Rauschen in der Zielfunktion sein.
- Nebenbedingungen, die sich aus dem Arbeitsbereich des Roboters, der Kollisionsvermeidung und der Mehrbildintegration ergeben, müssen geeignet berücksichtigt werden.

Formal läßt sich die Nebenbedingung, die sich aus dem Arbeitsbereich des Roboters ergibt, durch eine Funktion $C_{robot} : U \rightarrow \{0, 1\}$ beschreiben, welche für eine gegebene Aktion $u \in U$ angibt, ob der Roboter sie umsetzen kann.

Entsprechend kann das Kriterium zur Kollisionsvermeidung durch eine Funktion $C_{collision} : U \times I_k \rightarrow \{0, 1\}$ beschrieben werden, die basierend auf dem aktuellen Informationsstand $i_k \in I$ angibt, ob bei der Ausführung der Aktion $u \in U$ eine Kollision auftreten kann. Die zugrundeliegenden Überlegungen und der Zusammenhang mit dem eingesehenen Volumen sind Gegenstand von Abschnitt 4.3.3.

Die Integrierbarkeit des akquirierten Bildes zur evaluierten Aufnahmekonfiguration in den Bildverbund wird durch eine weitere Funktion $C_{regist} : U \times I_k \rightarrow \mathbb{R}$ beschrieben, die den prozentualen Anteil der Meßpunkte angibt, der bei der Ausführung der Aktion $u \in U$ erfaßt wird und bereits in einer der zuvor akquirierten Aufnahmen, die in $i_k \in I$ enthalten sind, ebenfalls erfaßt wurde.

Mit diesen Hilfsfunktionen kann die Optimierungsaufgabe nun formal beschrieben werden: gesucht ist in jedem Planungsschritt k diejenige Konfiguration u , die

$$\begin{aligned} \min G(u, i_k), \quad & C_{robot}(u) \neq 0 \quad \wedge \\ & C_{collision}(u, i_k) \neq 0 \quad \wedge \\ & C_{regist}(u, i_k) > \text{Schwellwert} \end{aligned} \quad (4.30)$$

für die in Gleichung 4.20 definierte Zielfunktion G erfüllt. Der in der Formel verwendete Schwellwert zur Integrierbarkeit wird in der Praxis auf 20% festgelegt.

Zur Lösung dieser Aufgabe wird eine leicht modifizierte uniforme Gridsuche angewendet, die insbesondere die geforderte optimierte Änderungsreihenfolge der Parameter unterstützt. Rauschen in der Zielfunktion spielt bei diesem Verfahren keine Rolle. Algorithmisch realisiert das Verfahren eine Reihenfolge von Parameterkombinationen, die in den Dimensionen des Suchraums jeweils die erforderliche Abtastdichte liefern. Das globale Minimum aller evaluierten Funktionswerte ist die berechnete Näherung der in Formel 4.30 definierten Aufgabenstellung.

4.6.2 Reihenfolge der Funktionsevaluationen

Wie in den vorangegangenen Abschnitten bereits deutlich wurde, stehen die frei wählbaren Parameter nicht gleichberechtigt nebeneinander, sondern das optische Zentrum des Aufnahmesystems nimmt eine Sonderstellung ein, die in der wechselseitigen Homöomorphismus-Beziehung der Bilddaten von ein und derselben Aufnahme position begründet liegt (Abschnitt 4.5.1). Zusätzlich kann auch in der Berechnung der Funktion $C_{collision}$ in vielen Fällen (einschließlich des in der vorliegenden Arbeit eingesetzten mobilen Meßsystems zur Innenraumrekonstruktion) eine rotationssymmetrische Form des

Effektors angenommen werden, so daß die Evaluation dieser Nebenbedingung ebenfalls für Parameterkombinationen mit identischer Aufnahmeposition nur einmal ausgeführt werden muß. Wir gehen also im Folgenden von einer vereinfachten Funktion $C'_{collision} : \mathbb{R}^3 \times I_k \rightarrow \{0, 1\}$ aus, die nicht von der vollständigen Aufnahmekonfiguration, sondern von deren Beobachtungspunkt abhängig ist.

Als algorithmische Beschreibung der Vorgehensweise bei der Reihenfolge der Parameteränderungen zur Abtastung des Parameterraums wird im Folgenden die Verschachtelung der relevanten Schleifen dargestellt:

Für alle evaluierten Beobachtungspunkte p mit $C'_{collision}(p, i_k) \neq 0$:

Berechne den sichtbaren Teil der teilvermessenen, ergänzten Oberfläche S .

Für alle Blickrichtungen $(\varphi_x, \varphi_y, \varphi_z)$ mit $C_{robot}(u) \neq 0$:

Für alle Sensor-Parameter $\{e_1, \dots, e_n\}$ mit $C_{regist}(u, i_k) > \text{Schwellwert}$:

Berechne $G(u, i_k)$

Speichere u zu $\min G(u, i_k)$

Durch die vorgeschlagene Struktur wird die Wiederverwertung von Zwischenergebnissen optimiert, da die rechenaufwändigeren Schritte jeweils in weiter außen liegenden Schleifen ausgeführt werden. Im folgenden Abschnitt werden Verfahren zur Abtastung des Parameterraums für jede Stufe der verschachtelten Schleifen-Struktur vorgeschlagen. Davon unabhängig erscheint es sinnvoll, die Menge der evaluierten Beobachtungspunkte über die Nebenbedingungen hinaus zunächst durch heuristische Verfahren auf besonders vielversprechende Teilvolumina zu begrenzen, beispielsweise auf Bereiche, in denen große, zusammenhängende unvermessene Randflächen des eingesehenen Volumens im Meßvolumen zu liegen kommen. Derartige Strategien wurden im Rahmen der vorliegenden Arbeit aber nicht umgesetzt und bleiben zukünftigen Arbeiten vorbehalten.

4.6.3 Abtasten des Parameterraums

Bei der Abtastung des Parameterraums im Rahmen der oben dargestellten algorithmischen Schleifen-Struktur können drei Fälle unterschieden werden. Für die Sensor-Parameter (intrinsische Parameter) sind typischerweise sinnvolle Wertebereiche bekannt, innerhalb derer jede Parameterkombination gültige Aufnahme-Konfigurationen ergibt. Für die extrinsischen Parameter gilt dies nicht, da große Bereiche des Parameterraums durch die Nebenbedingungen ausgeschlossen werden. Daher kann die Anzahl der tatsächlich zu evaluierenden Parametersätze aus einer vorgegebenen Abtastdichte nicht prädiziert werden. Folglich ist es wünschenswert, nicht die Abtastdichte, sondern den Rechenaufwand vorzugeben. Dies setzt eine Reihenfolge der Parameterevaluationen voraus, die nach Ablauf der zur Verfügung stehenden Zeit zu ei-

ner räumlich dichten gleichmäßigen Verteilung der evaluierten Punkte des Parameter-raums führt. Die hier informal verwendeten Begriffe der „Reihenfolge“ und der „räumlich dichten gleichmäßigen Verteilung“ werden dazu zunächst durch die entsprechenden formal definierten Begriffe ersetzt.

4.6.3.1 Definitionen

Wir unterscheiden eine Punktmenge von einer Punktfolge. Eine Punktfolge ist eine eindeutige Abbildung von \mathbb{N} in den \mathbb{R}^n . Eine Punktfolge heißt beschränkt, wenn ihr Wertevorrat beschränkt ist [12]. Für vollständig verwendete Punktfolgen ist die Unterscheidung von den Punktmengen zwar algorithmisch nicht relevant; die Reihenfolge der Punkte ist aber für die hier beabsichtigte unvollständige Verwendung von Punktfolgen entscheidend. Wegen der unbekannten Anzahl der benötigten Evaluationspunkte fokussiert sich die Betrachtung dabei auf unbeschränkte Punktfolgen.

Die wichtigste Eigenschaft der verwendeten unbeschränkten Punktfolge besteht darin, daß diese eine *dichte Punktmenge* ergibt. In einer dichten Punktmenge $P \subset \mathbb{R}^n$ existiert für jeden Punkt $x \in \mathbb{R}^n$ ein Punkt $p \in P$ für den $\|x - p\| < \varepsilon$ für ein beliebiges $\varepsilon \in \mathbb{R}$ gilt (die stärkere Forderung, daß jeder Punkt im Raum Element der unbeschränkten Punktfolge sein soll, kann nicht gestellt werden, da \mathbb{R}^n nicht abzählbar ist). Für den weiter unten vorgeschlagenen Ansatz sei dazu noch angemerkt, daß eine aus uniform verteilten Zufallspunkten gebildete unendlich große Punktmenge nur mit der Wahrscheinlichkeit 1 eine dichte Punktmenge ist, diese Eigenschaft aber nicht garantiert werden kann [57].

In der Praxis sind allerdings Punktmengen mit einer endlichen Anzahl an Elementen anzutreffen. Analog zum Begriff der *Auflösung* eines regulären Punktrasters definieren wir den Begriff der *Dispersion* für eine irreguläre Punktmenge P mit Hilfe des metrischen Raums X mit $P \subset X$, auf dem eine Norm $\|\cdot\|$ definiert ist, durch

$$\delta(P) := \sup_{x \in X} \min_{p \in P} \|x - p\| \quad (4.31)$$

Anschaulich wird die Definition⁸, wenn man sich am Beispiel der L_2 -Norm um jeden Punkt $p \in P \subset \mathbb{R}^3$ eine Kugel vorstellt, in der kein anderer Punkt $p' \in P$ liegt. Die Dispersion $\delta(P)$ bezeichnet dann den größten auftretenden Kugelradius; für die L_∞ -Norm analog die halbe Kantenlänge des größten auftretenden Würfels [57].

⁸Das Supremum G ist die kleinste obere Grenze einer Zahlenmenge X [12]. Falls X abgeschlossen ist, so entspricht das Supremum von X dem Maximum.

4.6.3.2 Punktfolgen in \mathbb{R}^n

Eine einfaches Verfahren zur Erzeugung einer unendlichen Punktfolge kann auf die von van der Corput [20, zitiert nach [57]] vorgeschlagene Zahlenfolge zurückgeführt werden. Im eindimensionalen Fall wird hierzu die binäre Darstellung der indizierenden Zahl $k \in \mathbb{N}$ in ihrer Ziffernreihenfolge umgekehrt und als Bruchteil des betrachteten Parameterraums interpretiert. Der jeweilige Zahlenwert $\varphi(k)$ ergibt sich durch

$$\varphi(k) = \sum_{i=0}^{\lfloor \lg k \rfloor} \frac{a_i}{2^{i+1}} \quad (4.32)$$

mit

$$k = \sum_{i=0}^{\lfloor \lg k \rfloor} a_i \cdot 2^i \quad (4.33)$$

und der ergänzenden Definition

$$\varphi(0) := 0 \quad (4.34)$$

Die entstehende Zahlenfolge $\{0, \frac{1}{2}, \frac{1}{4}, \frac{3}{4}, \frac{1}{8}, \frac{5}{8}, \frac{3}{8}, \frac{7}{8}, \dots\}$ erreicht nach 2^i Elementen für alle $i \in \mathbb{N}$ einen optimalen Dispersionswert im Verhältnis zur Anzahl der Elemente. Wird eine Punktfolge mit einer anderen Länge verwendet, so hat diese zumindest weitgehend gleichmäßige Dichte; insbesondere sind die zur 2^{i+1} -elementigen Folge noch fehlenden Elemente in der Folge der Länge $2^{i+1} - 2^j$ für $j \in \mathbb{N}$, $j < i$ gleichmäßig verteilt. Abbildung 4.9 zeigt die van-der-Corput-Folgen der Längen 1 bis 16.

Der Ansatz läßt sich auf einfache Weise auf Punktfolgen im \mathbb{R}^n erweitern, indem die Ziffernfolge der binären Darstellung nach der Umkehrung ihrer Reihenfolge in n -Tupel zusammengefaßt wird. Anschaulich betrachtet bestimmen dann jeweils n Ziffern hierarchisch einen Quadranten, Oktanten beziehungsweise den entsprechenden Halbraumschnitt der Dimension n . Ziffernweise betrachtet bilden die ersten Elemente aller vorliegenden n -Tupel die erste Koordinate, die zweiten Elemente die zweite Koordinate, und so weiter. Man erkennt aus dieser Struktur der Koordinatenerzeugung, daß in diesem Fall nach jeweils $(2^n)^i$ Elementen der Punktfolge für alle $i \in \mathbb{N}$ ein optimaler Dispersionswert im Verhältnis zur Anzahl der Elemente erreicht wird. Abbildung 4.10 zeigt die entstehende Punktfolge für \mathbb{R}^2 .

Im Rahmen der vorliegenden Arbeit wird die van-der-Corput-Punktfolge für \mathbb{R}^3 zur Erzeugung der Blickpunkte bei der numerischen Bestimmung des globalen Maximums der Zielfunktion eingesetzt. Auf diese Weise ist in der rechenintensivsten Stufe des in Abschnitt 4.6.2 beschriebenen Schemas eine dichte und gleichmäßige Abtastung garantiert. Dabei kann die Berechnung nach Ablauf eines vorgegebenen Zeitrahmens jederzeit abgebrochen werden.

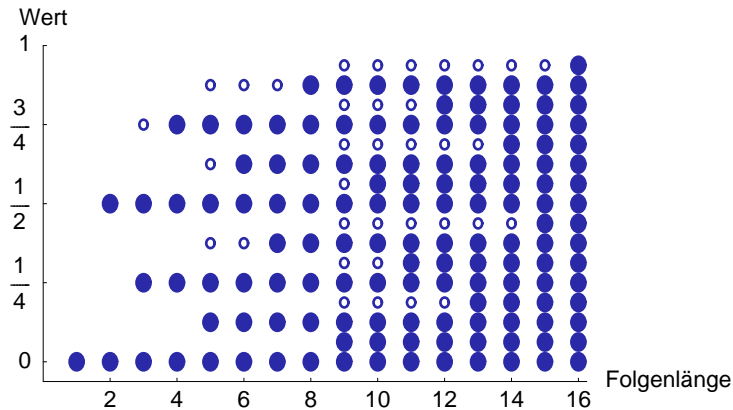


Abbildung 4.9: Die van-der-Corput-Folgen der Längen 1 bis 16 (gefüllte Kreise). Es wird deutlich, daß stets nach 2^i Elementen der Folge eine reguläre Anordnung mit minimaler Dispersion erzeugt wird. Die zur nächstlängeren 2^i -elementigen Folge noch fehlenden Elemente sind als leere Kreise dargestellt.

4.6.3.3 Rotationsfolgen in $SO(3)$

Die gleichmäßige, dichte Abtastung des Raums der dreidimensionalen Rotationen $SO(3)$ ist, wie die Darstellung der Parametrisierung in Abschnitt 4.3.2.1 bereits vermuten läßt, mit einfachen Ansätzen (wie etwa dem gleichmäßigen Abtasten des \mathbb{R}^3 der Euler-Winkel) nicht zu erreichen. Über die topologische Äquivalenz läßt sich zeigen, daß vielmehr eine Gleichverteilung der Abtastpunkte auf der in den vierdimensionalen Raum eingebetteten dreidimensionalen Kugeloberfläche des \mathbb{S}^3 erreicht werden muß, welche durch die korrespondierenden Quaternionen gebildet wird (die Mehrdeutigkeit der Zuordnung von $SO(3)$ nach \mathbb{S}^3 vereinfacht hier die Definition, da die Doppeldeutigkeit eine Unterscheidung der Hemisphären überflüssig macht).

Wegen des geringeren Evaluationsaufwands der Zielfunktion bei alleiniger Änderung der Rotationsparameter im Rahmen des in Abschnitt 4.6.2 beschriebenen Schemas erscheint es möglich, eine Rotationsfolge basierend auf Zufallszahlen zu verwenden. Die Problematik der nicht garantierten Dispersionsbegrenzung der Abtastung kann so auf statistische Weise durch eine entsprechend große Länge der verwendeten Rotationsfolge in ihrer Wahrscheinlichkeit bis auf ein in der Praxis akzeptables Maß reduziert werden.

Ein einfaches Verfahren zur Erzeugung von Rotationsfolgen wurde von Arvo [5] vorgeschlagen. Es werden drei unabhängige, gleichverteilte Zufallszahlen $z_1, z_2, z_3 \in (0, 1]$ benötigt. Aus diesen kann ein gleichverteiltes Zufallsquaternion q durch

$$q = (\sqrt{1 - z_1} \sin 2\pi z_2, \sqrt{1 - z_1} \cos 2\pi z_2, \sqrt{z_1} \sin 2\pi z_3, \sqrt{z_1} \cos 2\pi z_3) \quad (4.35)$$

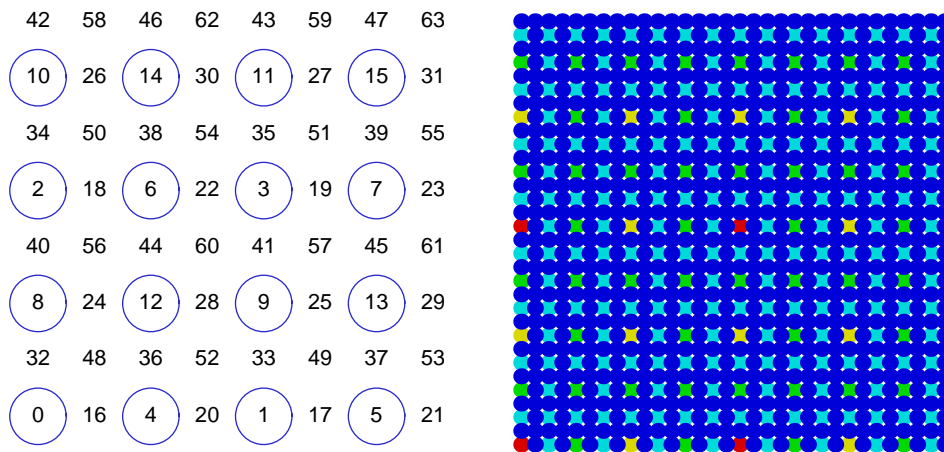


Abbildung 4.10: Die ersten 64 Punkte (links) bzw. die ersten 1024 Punkte (rechts) der van-der-Corput-Punktfolge in 2D. Rechts werden die ersten Punkte rot, die letzten blau dargestellt. Es wird deutlich, daß stets nach 4^i Elementen der Punktfolge ein reguläres Gitter mit minimaler Dispersion erzeugt wird.

erzeugt werden. Auf den Beweis der Gleichverteilung wird an dieser Stelle verzichtet; Details zur Herleitung finden sich neben der Originalarbeit auch in [54]. Die Gleichverteilung der erzeugten Rotationen wird in der Visualisierung in Abbildung 4.11 im Vergleich mit einfachen, eulerwinkelbasierten Ansätzen deutlich. Die Visualisierung zeigt andererseits auch, daß bereits die Positionen der rotierten tangentialen Geradenabschnitte auf der Kugeloberfläche in sehr unterschiedlichem Abstand zueinander liegen; betrachtet man auch die visuell weniger gut erkennbare dritte Rotation um die Normale der Kugeloberfläche, so wird deutlich, daß das Verfahren den Raum der 3D-Rotationen relativ ineffizient abtastet.

Eine Alternative bildet die Erzeugung einer deterministischen, unbeschränkten und dichten Rotationsfolge. Yershova et al. [117] schlagen ein derartiges Verfahren auf Basis platonischer Polytope vor. Sie definieren ein Raster auf den Elementen des jeweiligen Polytops und projizieren dieses Raster anschließend auf die Kugeloberfläche \mathbb{S}^3 . Hierzu stehen im \mathbb{R}^4 6 reguläre (platonische) Polytope zur Verfügung, die aus 5, 8, 16, 24, 120 beziehungsweise 600 Polyedern aufgebaut sind. Der Implementierungsaufwand steigt dabei mit der Komplexität des Polytops, und das Verständnis und die Validierung werden durch die Dimensionalität erschwert. Daher wird im Folgenden eine neue, vereinfachte Variante des gleichen Ansatzes vorgeschlagen, deren Konzepte mit drei Dimensionen darstellbar und implementierbar sind.

Die Rückführung auf dreidimensionale Strukturen wird plausibel, wenn man sich die Definition eines Quaternions durch die Drehachse w mit $\|w\| = 1$ und den Drehwinkel ϑ nach Gleichung 4.9 in Erinnerung ruft. Dabei wird die Drehachse zwar mit $\sin \frac{\vartheta}{2}$ skaliert; ihre Richtung bleibt aber vom Drehwinkel unbeeinflusst. Daher können w und ϑ

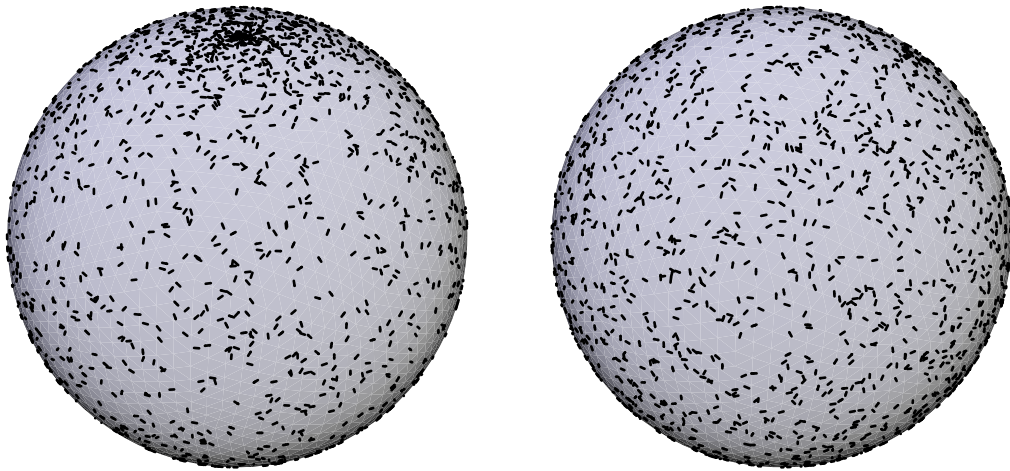


Abbildung 4.11: Naive (links, gleichverteilte Euler-Winkel) und gleichverteilte Abtastung (rechts, Quaternionen-Ansatz nach Gleichung 4.35) von $SO(3)$. Dargestellt ist ein tangentialer Geradenabschnitt auf einer Kugel um den Ursprung, welcher durch die jeweils 3078 erzeugten Rotationen transformiert wurde.

getrennt bestimmt werden; wegen $\|w\| = 1$ liegen die dreidimensionalen Imaginärteile eines Quaternions stets in \mathbb{S}^2 , also auf der Oberfläche der Einheitskugel in \mathbb{R}^3 . Der skalare Realteil kann unabhängig gewählt werden.

Im Kontext der vorliegenden Arbeit erscheint es dabei sehr sinnvoll, den Real- und den Imaginärteil des Quaternions nicht nur unabhängig zu berechnen, sondern dabei auch die Dichte unterschiedlich zu wählen. Geometrisch entspricht der Imaginärteil, also die Drehachse, bei entsprechender Verwendung der erzeugten Rotation der Hauptachse der Kamera, also der Blickrichtung. Hier ist, insbesondere bei der externen Digitalisierung kleiner Objekte, eine hohe Auflösung erforderlich, da bereits kleine Änderungen der Blickrichtung dazu führen, daß das Objekt nicht mehr in das Sichtfeld des Sensors fällt. Ein entsprechend großer Anteil der erzeugten Drehachsen führt daher zu Quaternionen beziehungsweise Aufnahmekonfigurationen, die die Nebenbedingungen nicht erfüllen. Der Realteil entspricht dagegen geometrisch der Drehung um die Kamerahauptachse. Hier ist typischerweise eine niedrigere Dichte sinnvoll⁹.

Ziel des hier vorgeschlagenen Verfahrens ist also die möglichst verzerrungsfreie Abbildung einer zweidimensionalen Punktfolge auf die Oberfläche der Einheitskugel, um

⁹Die Anforderung, unterschiedliche Dichten in Real- und Imaginärteil des Quaternions beim Abtasten des $SO(3)$ zu berücksichtigen, tritt in den Originalarbeiten [117, 54] nicht auf, da die beabsichtigte Anwendung dort im Bereich *path planning* liegt.

auf diese Weise die Imaginärteile der Quaternionen zu bestimmen. Analog zur Vorgehensweise von Yershova et al. [117] wird die Punktfolge zunächst auf den planaren Flächenelementen eines platonischen Polyeders erzeugt. Wegen der einfachen Umsetzung wird hierzu der Würfel ausgewählt, der punktsymmetrisch zum Ursprung im Raum liegt. Die Elemente der Punktfolge auf der Würfel­fläche können nun durch Normalisierung auf die Einheitskugel abgebildet werden.

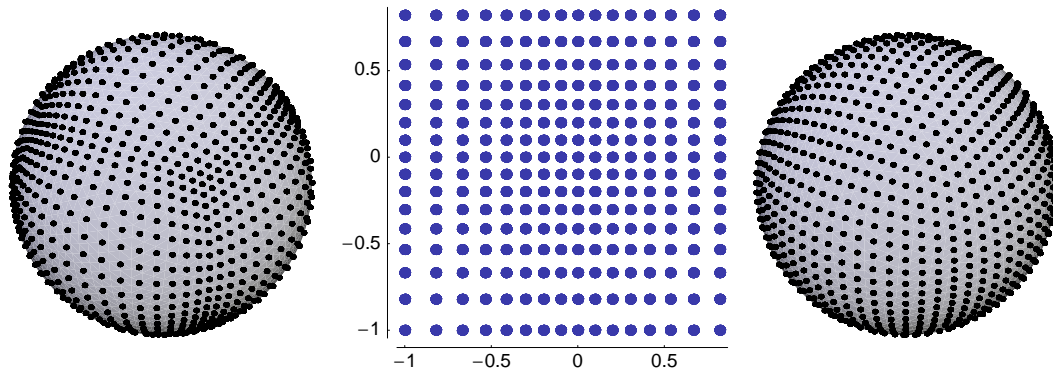


Abbildung 4.12: Ein gleichmäßiges Raster mit je 256 Punkten auf den Flächen eines Würfels erscheint nach der Projektion auf die Kugel zunächst verzerrt (links). Eine Vorverzerrung des Rasters mit der in Gleichung 4.36 definierten Abbildungsfunktion (mitte) beseitigt die Verzerrung auf der Kugeloberfläche weitgehend (rechts).

Wie in Abbildung 4.12 zu erkennen ist es sinnvoll, die Elemente der Punktfolge auf der Würfeloberfläche verzerrt zu definieren, um nach der Projektion auf die Einheitskugel einen gleichmäßigen Winkelabstand zwischen den Punkten zu erhalten. Die entsprechende Korrekturabbildung ist aufgrund der einfachen Geometrie des Würfels mit den Eckpunkten $(\pm 1, \pm 1, \pm 1)$ durch eine ebenfalls einfache Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ mit

$$f(u, v) = \begin{pmatrix} -\cot\left(\frac{\pi}{4}(u+2)\right) \\ -\cot\left(\frac{\pi}{4}(v+2)\right) \end{pmatrix} \quad (4.36)$$

zu realisieren. Für den Erhalt der Dichte der Punktfolge bei der Anwendung der Korrekturabbildung sind dabei Stetigkeit und Reproduktion der Ränder entscheidend. Die Kotangensfunktion $\cot x$ ist für alle $x \neq i \cdot \pi$ mit $i \in \mathbb{N}$ stetig, also auch im hier genutzten Bereich $(\frac{\pi}{4}, \frac{3\pi}{4})$. Ferner ist wegen $\cot \frac{\pi}{4} = 1$ und $\cot \frac{3\pi}{4} = -1$ auch die Randreproduktion erfüllt. Abbildung 4.12 zeigt ein korrigiertes Raster und die Wirkung auf der Kugeloberfläche.

Mit diesem Ansatz lassen sich dichte unbeschränkte Rotationsfolgen aufbauen, indem zunächst die entsprechende dreidimensionale van-der-Corput-Punktfolge erzeugt

wird. Die ersten beiden Koordinaten jedes Elements der 3D-Punktfolge definieren dabei sechs Punkte auf den quadratischen Würfelflächen, aus denen sich durch die in Gleichung 4.36 und Normalisierung die Drehachsen der entsprechenden Quaternionen berechnen lassen. Die dritte Koordinate jedes Elements der 3D-Punktfolge definiert für diese sechs Quaternionen jeweils den Drehwinkel. Im Mittel ergibt sich ein Verhältnis von ungefähr 1 : 4.34 zwischen der Winkelauflösung des Real- und des Imaginärteils des Quaternions in der Rotationsfolge.

Für Systeme mit größerem Meßabstand kann dieses Verhältnis der Winkelauflösungen stufenweise angepaßt werden, indem die Würfelflächen in $n \times n$ Teilquadrate zerlegt werden, so daß jedes Element der 3D-Punktfolge $6n^2$ Quaternionen erzeugt. Abbildung 4.13 visualisiert die so erzeugte Rotationsfolge in Analogie zur Visualisierung

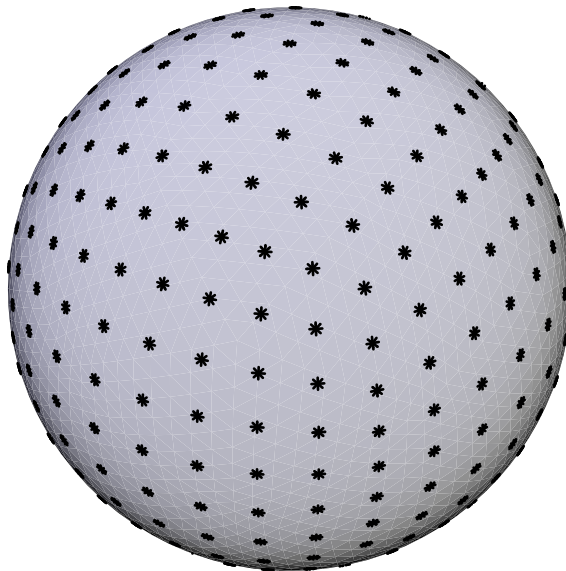


Abbildung 4.13: Deterministische Abtastung von $SO(3)$ durch eine unbeschränkte dichte Rotationsfolge basierend auf der van-der-Corput-Punktfolge. Dargestellt ist ein tangentialer Geradenabschnitt auf einer Kugel um den Ursprung, welcher durch die jeweils 3078 erzeugten Rotationen transformiert wurde (vergleichbar der Darstellung in Abb. 4.11).

der Zufallsrotationsfolgen in Abbildung 4.11. Die Rotationsfolge erreicht stets nach $6n^2 \cdot 8^i$ Elementen minimale Dispersionswerte im Verhältnis zur Anzahl der Elemente.

Im Rahmen der vorliegenden Arbeit wird die auf Basis der van-der-Corput-Punktfolge definierte Rotationsfolge zur Erzeugung der extrinsischen Rotationsparameter

bei der numerischen Bestimmung des globalen Maximums der Zielfunktion eingesetzt. Auch hier kann die erzeugte Rotationsfolge an beliebiger Stelle nach Ablauf eines vorgegebenen Zeitrahmens abgebrochen werden. Dabei ist jederzeit eine dichte und gleichmäßige Abtastung des Raums der 3D-Rotationen $SO(3)$ garantiert.

4.7 Abbruchkriterium des Gesamtverfahrens

Die Wahl des Abbruchkriteriums ist entscheidend für die Qualität iterativer Verfahren. Erfolgt der Abbruch zu früh, so wird eine suboptimale Lösung geliefert, erfolgt er zu spät, so wächst die Größe des entstehenden Bildverbundes, ohne zusätzliche Informationen zu liefern, und erschwert so die Weiterverarbeitung.

Es werden zwei Arten von Abbruchkriterien kombiniert, von denen der erste die Konvergenz des Verfahrens betrifft. Das Konvergenzkriterium ist der Absolutwert des prädizierten Qualitätsgewinns der optimierten nächsten Aufnahmekonfiguration. Sinkt dieser unter einen Schwellwert, so wird das Gesamtverfahren beendet. Das zweite Kriterium betrifft Aufnahmekonfigurationen, bei denen der erwartete Qualitätsgewinn nicht eintritt, weil beispielsweise eine unkooperative Oberfläche eingesehen wird, zu der keine Meßdaten erhoben werden können. Hier ist eine manuelle Verbesserung der Vermeßbarkeit der betreffenden Oberfläche nötig. Alternativ kann der betreffende Bereich von der weiteren Optimierung ausgeschlossen werden, indem den betreffenden Grenzflächen manuell die Zielqualität als Oberflächenattribut zugewiesen wird.

Aufgrund dieser Überlegung kann für den Fall unkooperativer Oberflächen als Abbruchkriterium die erreichte Qualitätssteigerung durch die durchgeführte Aufnahme herangezogen werden. Durch Auffinden des Minimums der in Gleichung 4.20 definierten Zielfunktion G kann mit

$$u_{k+1} = \min_u G(u, i_k) + \Delta u_{k+1} \quad (4.37)$$

die geplante und mit der Abweichung Δu_{k+1} tatsächlich ausgeführte Aktion u_{k+1} bestimmt werden. Analog zu der in Gleichung 4.16 definierten Funktion β_{k+1} , die auf Basis des Informationsstandes in Iteration k prospektiv die erwartete Qualität beschreibt, wird hier eine Funktion

$$\bar{\beta}_{k-1} : \mathbb{R}^3 \times U \times I_k \quad (4.38)$$

definiert, die retrospektiv auf Basis des Informationsstandes $i_k \in I_k$ der aktuellen Iteration k einem Punkt $\mathbf{p} \in S_k \subset \mathbb{R}^3$ den Qualitätswert $\bar{\beta}_{k-1}(\mathbf{p}, u, i_k)$ zuweist, der vor Integration der im Rahmen von Aktion u akquirierten Aufnahme vorlag.

Mit Hilfe dieser Funktion läßt sich eine Testgröße t berechnen, indem die Qualitätssteigerung über dem Teil der Oberfläche integriert wird, der bei der durchgeführten

Aktion u_{k+1} vom Sensor eingesehen werden konnte. Formal lässt sich dies durch

$$t_k = \frac{\int \int_{\mathbf{p} \in S_k} w_{\text{sichtbar}}(\mathbf{p}, u_k, i_k) \cdot \min\{\beta_{\max}(\mathbf{p}), \bar{\beta}_{k-1}(\mathbf{p}, u, i_k)\} dS}{\int \int_{\mathbf{p} \in S_k} w_{\text{sichtbar}}(\mathbf{p}, u_k, i_k) \cdot \min\{\beta_{\max}(\mathbf{p}), \tilde{\beta}_k(\mathbf{p}, i_k)\} dS} \quad (4.39)$$

ausdrücken. Trifft nach Integration der Aufnahme zur Aktion u_{k+1} das Abbruchkriterium $t_{k+1} > t_{\text{Schwellwert}}$ zu, so terminiert das Verfahren, da aufgrund der nahezu unveränderten Datensituation auch in weiteren Iterationen keine Verbesserung erreicht werden kann.

4.8 Zusammenfassung eigener Ergebnisse

In diesem Kapitel wurde zunächst das Problem der Aufnahmeplanung formal spezifiziert und auf die in der Planungstheorie etablierten Begriffe von Umgebung, Zustand beziehungsweise Informationsstand und Aktion zurückgeführt. Dabei wurde besonders auf die Struktur des Aktionsraums eingegangen, da dieser den schwierig zu handhabenden Raum der 3D-Rotationen als Teilraum enthält.

Einer der beiden Kernpunkte des Kapitels ist die vorgeschlagene Zielfunktion, die auf Basis der im vorangegangenen Kapitel 3 definierten Grenzfläche des eingesehenen Volumens jeder Aufnahmekonfigurationen eine Bewertung zuordnet. Über die aus der Literatur bekannten Strategien hinaus berücksichtigt die in der vorliegenden Arbeit vorgeschlagene Zielfunktion auch den Aufwand der zusätzlichen Aufnahme (im Sinne einer „Kosten-Nutzen-Rechnung“) sowie die bereits erreichte und die gewünschte Qualität der Oberflächenvermessung.

Der zweite Kernpunkt des Kapitels ist die effiziente Berechnung der Zielfunktion. Hier wurden neuartige Verfahren vorgeschlagen, mit denen die Zielfunktion in der wählbaren Genauigkeit eines Bildrasters berechnet werden kann. Dazu wurden die wesentlichen Teilschritte der Berechnung systematisch auf computergraphische Basisfunktionen zurückgeführt, die sich mit Hilfe spezialisierter Graphik-Hardware beschleunigt berechnen lassen. Wegen der zentralen Bedeutung der Sichtbarkeitsrechnung wird dabei zwischen unifokalen Sensoren, die nur einen Beobachtungspunkt nutzen, und triangulationsbasierten Sensoren mit zwei Beobachtungspunkten unterschieden. Letztere erfordern Verfahren der Schattenrechnung, für die ein neuartiges Verfahren zur computergraphischen Modellierung des Sensor-Verhaltens vorgeschlagen wurde.

Einen weiteren Beitrag bildet die Organisation der globalen Optimierung der Zielfunktion. Hier wurde ein Verfahren vorgeschlagen, das unter Berücksichtigung des abzutastenden Raums der 3D-Rotationen ($SO(3)$) eine Evaluationsreihenfolge nutzt, die

jederzeit abgebrochen werden kann und dennoch eine gleichmäßige Abtastastung des Parameterraums garantiert.

Das Abbruchkriterium bildet schließlich die letzte Voraussetzung zum automatischen Ablauf des vorgeschlagenen Gesamtverfahrens.

KAPITEL 5

Bewertung und Ergebnisse

5.1 Einleitung

Die vorliegende Implementierung der in den Kapiteln 3 und 4 vorgeschlagenen Verfahren wirft die Frage auf, wie diese bewertet und mit anderen Verfahren auf objektive Weise verglichen werden kann. Die Antwort ergibt sich zum einen aus der Prüfung der generellen Funktionsfähigkeit des Verfahrens: Wird die Vermessung eines unbekannten Objekts oder einer unbekannten Umgebung tatsächlich so geplant, daß der zyklische Gesamt Ablauf der Bildakquisition und -integration störungsfrei ablaufen kann und im Ergebnis zu einem brauchbaren Modell führt? Zum anderen müssen die Fragen nach der Ergebnisqualität, nach dem Ressourcenverbrauch und nach der Planungseffizienz durch geeignete Maßzahlen beantwortet werden, die im folgenden Abschnitt vorgeschlagen werden. Beispiele eigener Ergebnisse werden in Abschnitt 5.3 vorgestellt.

5.2 Evaluationskriterien

5.2.1 Überblick eigener Vorschläge

Die Planung der Bildakquisition verfolgt verschiedene, teilweise konkurrierende Ziele. Daher ist es nicht sinnvoll, ein einziges Kriterium zur Bewertung der geplanten Aufnahmesequenz heranzuziehen. Beispielsweise ist es für die Planung vorteilhaft, zu einem möglichst frühen Zeitpunkt Aufschluß über die ungefähre Objekt- beziehungsweise Umgebungsgeometrie zu erhalten, da auf diese Weise größere Freiheit bei der

Auswahl der optimalen Beobachtungspunkte unter Ausschluß von Kollisionen und Eigenverdeckungen gegeben ist. Dieses Ziel steht aber der angestrebten Erfassung der Oberflächen mit einer vorgegebenen Abtastdichte zunächst entgegen.

Aus den in Abschnitt 4.1 formulierten Zielen der Aufnahmeplanung können zunächst die vier folgenden Kriterien abgeleitet werden, die die Qualität der Planung kennzeichnen:

1. Der **Volumenanteil**, der vom Meßsystem eingesehen und als leer identifiziert werden konnte, relativ zum gesamten Leervolumen. Die Maximierung dieses Wertes erleichtert insbesondere bei der Erfassung der Umgebung mit mobilen Robotern die Planung, da er mit der Bewegungsfreiheit des Aufnahmesystems korreliert.
2. Der Anteil der mit einer gegebenen Mindestzuverlässigkeit **vermessenen Objektoberfläche** relativ zur gesamten Objektoberfläche.
3. Die erreichte **Abtastdichte** an jedem Punkt der Oberfläche und, damit verwandt, die **Qualität** der jeweiligen Meßpunkte.
4. Das Datenvolumen und der **Ressourcenverbrauch**, also insbesondere die Anzahl der Akquisitionsschritte, die Anzahl der Meßpunkte (sofern variabel) und (mit Einschränkungen bei der Vergleichbarkeit) der Zeitaufwand.

Da das vierte Kriterium ein Kostenkriterium ist, liegt es nahe, die übrigen Kriterien als erreichten Gewinn im Vergleich zum Aufwand darzustellen. Die Evaluierung der in der vorliegenden Arbeit gezeigten Ergebnisse konzentriert sich auf die ersten beiden Kriterien. Die vorgeschlagenen Kriterien und ihre Berechnung werden in den folgenden Abschnitten beschrieben.

Die Anwendung der Kriterien erfordert bekannte Referenzwerte, insbesondere eine exakte und vollständige Beschreibung der zu rekonstruierenden Oberflächen. In Versuchen mit künstlichen Oberflächenmodellen ist dies einfach zu erreichen. In der vorliegenden Arbeit wurde ein Akquisitions-Simulator eingesetzt, der die geplanten Tiefenbild-Aufnahmen aus einem Dreiecksnetz-Modell erzeugt. Reale Szenen sind jedoch oftmals nicht mit vertretbarem Aufwand vollständig zu erfassen. Ersatzweise wird in diesem Fall der am Ende des Rekonstruktionsprozesses generierte Datensatz als Referenz zur Evaluierung des Planungsverlaufs eingesetzt.

5.2.2 Volumetrisches Kriterium

Die volumetrische Betrachtung der Rekonstruktionssituation während des Ablaufs der Planung gründet sich auf die bereits in Kapitel 3 thematisierte Erkenntnis, daß sich zwischen einem auf der Oberfläche erfaßten Meßpunkt und der zugehörigen Beobachtungsposition kein verdeckendes Hindernis befinden kann. Dieser Zusammenhang

ist nicht nur für die konsistente Definition der teilvermessenen Objektoberfläche nützlich, sondern impliziert auch direkt den für die Planung verfügbaren Bewegungsraum des Robotersystems, in welchem Kollisionen sicher ausgeschlossen werden können.

Die Implementierung dieses Kriteriums basiert auf der Diskretisierung des betrachteten Raums mit Hilfe eines Voxel-Grids, in welchem für jeden Voxel die Sichtbarkeit bezüglich der evaluierten Menge der Einzelaufnahmen bestimmt wird. Sichtbar ist ein Voxel entsprechend Gleichungen 3.3 und 3.4 dann, wenn er in eines der Bilder projiziert zwischen der im Bild repräsentierten Oberfläche und der Beobachungsposition zu liegen kommt. Als Maßzahl wird daraus der Volumenanteil abgeleitet, der zu einem evaluierten Zeitpunkt innerhalb des Planungsablaufs als leer klassifiziert werden kann.

Neben dem relativen Freiraum bei der Planung der Roboterbewegung erlaubt das volumetrische Kriterium auch eine Aussage über die noch nicht aufgelösten *void*-Volumen. Die Planung wird rechnerisch erheblich vereinfacht, wenn statt der unvermessenen Grenzfläche eines noch nicht eingesehenen Teilvolumens eine beliebig schlecht vermessene Oberfläche berücksichtigt werden kann, da auf diese Weise eine verlässliche quantitative Aussage über den Gewinn einer Bildakquisition des betroffenen Oberflächenbereichs in die Planung einbezogen werden kann.

Für die externe Rekonstruktion von Gegenständen ist das volumetrische Kriterium nicht geeignet, da es im Unterschied zur internen Rekonstruktion von Innenräumen keine durch die Meßaufgabe vorgegebene Grenze des zu betrachtenden Volumens gibt. In der Praxis wird der zu betrachtende Bereich in Abhängigkeit von der *a priori* verfügbaren Kenntnis über die Objektgröße möglichst klein gewählt. Dabei hat beispielsweise der Arbeitsabstand des Meßsystems zur Objektoberfläche Einfluß auf die einzuhaltenden Sicherheitsabstände und damit auf die Größe des Volumens, das für eine sichere Vermessung als leer bekannt sein muß.

5.2.3 Abtastdichte

Alle im Rahmen der vorliegenden Arbeit angewendeten 3D-Meßsysteme liefern die gemessenen Datenpunkte in einem 2D-Bildraaster, welches die Messung der erreichten Abtastdichte auf einfache Weise ermöglicht. Bevor diese beschrieben wird, soll jedoch zunächst auf den allgemeinen Fall eingegangen werden, in dem das 3D-Meßsystem lediglich eine unstrukturierte Punktwolke als Ergebnis liefert. Die vorgeschlagene Lösung für diese Datensituation zeigt die prinzipielle Vergleichbarkeit der Evaluationsergebnisse mit anderen Verfahren zur Aufnahmeplanung, die nicht auf bildartigen Strukturen aufsetzen.

5.2.3.1 Unstrukturierte Punktwolken

Eine in einer unstrukturierten Punktwolke enthaltene Einzelmessung bezieht sich auf ein unendlich kleines Flächenstück. Daher muß zunächst eine geeignete Methode gefunden werden, um die Dichte dieser Einzelmessungen auf der rekonstruierten Oberfläche abzuschätzen. Hierzu können etablierte Methoden der Statistik herangezogen werden, wie sie zum Beispiel Silverman [100] beschreibt. Die übliche Vorgehensweise ersetzt die nativ vorliegenden δ -Funktionen (eine endliche Menge Information, nämlich eine Messung, auf unendlich kleiner Fläche, dem 3D-Punkt) durch glatte *kernel*-Funktionen $k : \mathbb{R} \rightarrow \mathbb{R}$, die jedem Oberflächenpunkt x in Abhängigkeit vom Abstand zum Meßpunkt einen durch diesen verursachten Dichtebeitrag $k(|x - x_i|)$ zuordnen. Dabei ist das Integral der *kernel*-Funktion 1. Die lokale Dichte kann auf diese Weise an jedem Punkt der rekonstruierten Oberfläche durch Summation der Dichtebeiträge geschätzt werden (Abbildung 5.1).

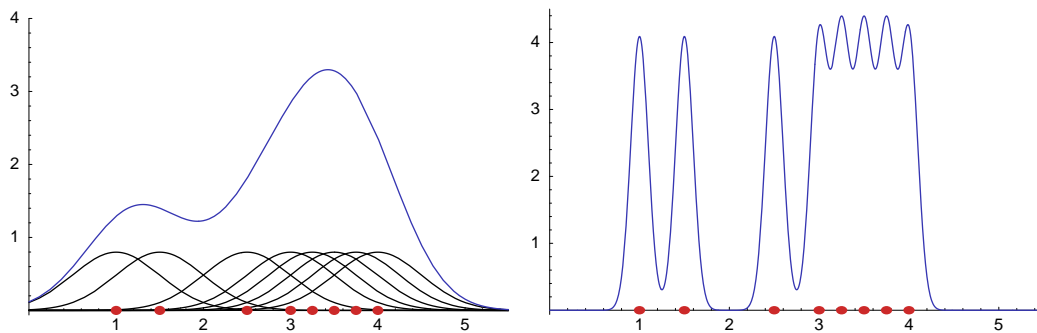


Abbildung 5.1: Dichteabschätzung mit einzelnen *kernel*-Funktionen (links) und Dichteabschätzung unter Verwendung von schmalbandigen kernels.

Der entscheidende Aspekt dieser Abschätzung ist die Wahl der Bandbreite des *kernels*, die in der Literatur als wesentlich wichtiger eingeschätzt wird als die Wahl der *kernel*-Funktion selbst [100]. Dabei ist eine datenunabhängige Wahl der Bandbreite sinnvoll, damit unvermessenen Bereichen der rekonstruierten Oberfläche auch tatsächlich die Abtastdichte 0 zugeordnet wird. Gleichzeitig ist durch die vorgegebene gewünschte Abtastdichte eine sinnvolle Maßzahl verfügbar, aus der die Bandbreite abgeleitet werden kann. Aus implementierungstechnischer Sicht ist es darüberhinaus sinnvoll, die Bandbreite so klein zu wählen, daß die euklidische 3D-Distanz als Distanz-Metrik verwendet werden kann.

Durch die Berücksichtigung aller Meßpunkte innerhalb der Einflußregion der *kernel*-Funktion ist das Verfahren rechnerisch aufwändig und sollte lediglich als Notlösung eingesetzt werden, wenn die ursprüngliche Rasterung der Meßdaten nicht mehr rekonstruiert werden kann. Eine mögliche Vereinfachung ergibt sich, wenn die Meßpunktdichte abschnittsweise auf der Oberfläche ermittelt wird. Dies ist zum Beispiel durch

Projektion der Meßpunkte auf ein Dreiecksnetz einfach möglich, indem die Dreiecksvermaschung als Abschnittsaufteilung gewählt wird. Für die weitergehende Analyse zum Beispiel mit Schwellwerten ist allerdings ein sinnvolles Verhältnis von Meßpunkten pro Abschnitt eine wichtige Grundvoraussetzung. Im Rahmen der vorliegenden Arbeit wurde das Verfahren nicht praktisch angewendet.

5.2.3.2 Tiefenbilder

Mit der in Abschnitt 3.3 eingeführten Projektionsfunktion h , die beliebige 3D-Punkte auf das Bilddraster der Tiefenbilder abbildet, kann auch die erreichte Abtastdichte gemessen werden. Die Details dieser Berechnung für einzelne Punkte werden in Abschnitt 3.4.4.1 in den Gleichungen 3.25 und 3.26 beschrieben.

Der algorithmische Ablauf der Dichteberechnung aus mehreren Tiefenbildern gestaltet sich folgendermaßen:

1. Projiziere den Punkt $\mathbf{p} \in \mathbb{R}^3$ in jedes Tiefenbild i unter Verwendung der externen Orientierung (R_i, t_i) und der Projektionsfunktion h .
2. Vergleiche den an der Stelle $h(R_i \cdot \mathbf{p} + t_i)$ gespeicherten Tiefenwert (Sichtbarkeitsprüfung).
3. Berechne die Abtastdichte mit Gleichung 3.26; falls $h(R_i \cdot \mathbf{p} + t_i)$ verdeckt oder außerhalb der Bildgrenzen liegt, sei die Abtastdichte 0.
4. Kombiniere die mit Schritt 3 berechneten Abtastdichten entsprechend dem eingesetzten Auswerteverfahren (typischerweise Summierung oder Maximumbildung).

Da an dieser Stelle keine Subpixel-Genauigkeit erforderlich ist, kann die Abtastdichte alternativ auch durch Berechnung der Fläche des 2×2 -patches, in das der projizierte Punkt fällt, berechnet werden. Dies wird in den meisten Fällen rechenaufwändiger sein und ist daher vor allem dann sinnvoll, wenn die Projektionsfunktion und ihre partiellen Ableitungen nicht bekannt sind, so daß man sich zum Beispiel mit Suchverfahren behelfen muß.

5.3 Ergebnisse und Diskussion

5.3.1 Interne Rekonstruktion

Die praktische Einsatzfähigkeit des Verfahren für Innenräume wird anhand zweier komplexer Beispielszenen demonstriert. Die erste Beispielszene ist das in Abbildung 5.2

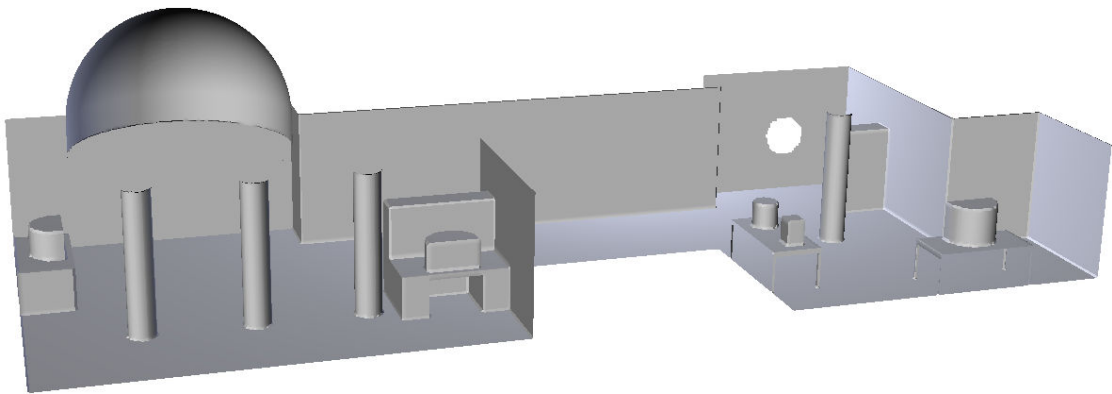


Abbildung 5.2: Gerenderte Ansicht der künstlichen Testszene unter Ausblendung rückwärts eingesehener Oberflächen (backface culling).

gezeigte künstliches Modell von $20\text{m} \times 7\text{m} \times 5.3\text{m}$, das im Zusammenspiel mit einem Bildakquisitions-Simulator eingesetzt wurde. Diese Vorgehensweise hat den Vorteil eines geringen Ressourcenverbrauchs, da das Verfahren unüberwacht ablaufen kann. Gleichzeitig können alle relevanten Eigenschaften eines realen mobilen Akquisitionssystems (Bildgeometrie, Rauschen, Meßbereich) nachgebildet werden. Die zweite Beispielszene ist eine reale typische Labor-Umgebung von $10\text{m} \times 6\text{m} \times 4.3\text{m}$. Diese für die Beurteilung des Verfahrens wichtige Situation kann nur an einem einzigen Beispiel gezeigt werden, da der hohe Aufwand¹ einer Wiederholung entgegensteht. Der wesentliche Unterschied zwischen der künstlichen und der realen Szene besteht in dem sehr viel höheren Detailreichtum, vor allem durch Kabel und herumliegende Gegenstände.

5.3.1.1 Künstliche Testszene

Die künstliche Testszene wurde im Rahmen der vorliegenden Arbeit eigens modelliert und enthält eine Reihe herausfordernder Details: in einer Struktur von zwei durch einen Korridor verbundenen Räumen befinden sich vier Säulen, zwei halbhohe Wände sowie je zwei arbeitstischartige Geometrien, die Tischplatten auf Schubladenelementen beziehungsweise auf Tischbeinen mit daraufstehenden Bildschirmen nachbilden. Um auch das Verhalten des Verfahrens an größeren nichtplanaren Flächen zu simulieren enthält die Szene darüberhinaus eine halbkugelförmige Kuppel. Im Ergebnis führt die Anordnung der modellierten Teilgeometrien dazu, daß aus nahezu jedem Blickpunkt

¹Es handelt sich um eine mehrstündige Messung außerhalb der normalen Arbeitszeiten, da der Raum während der Messung wegen des zur Lichtlaufzeitmessung eingesetzten Lasers nicht betreten werden darf. Zusätzlich ist aufwändige Vorbereitung erforderlich, da alle spiegelnden und transparenten Flächen (Bildschirme, Fenster) sinnvollerweise vorher abgeklebt werden müssen.

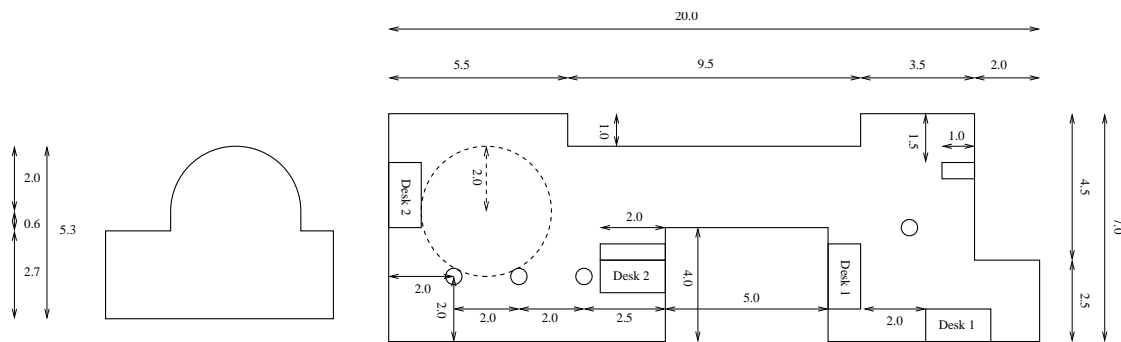


Abbildung 5.3: Zeichnung zur künstlichen Testszene in Seitenansicht (links) und Draufsicht (rechts).

eine große Zahl von Eigenverdeckungen auftritt. Gleichzeitig sind einige der Strukturen im Durchmesser in der Größe der Sensorauflösung (5 cm dicke „Tischbeine“) ausgeführt, und die Szene enthält einige „Löcher“, die unvermessbare Teilbereiche simulieren. Auf diese Weise werden auch Meßfehler und Meßpunktausfälle in die Prädiktion eingetragen, mit denen die Robustheit des Verfahrens untersucht werden kann.

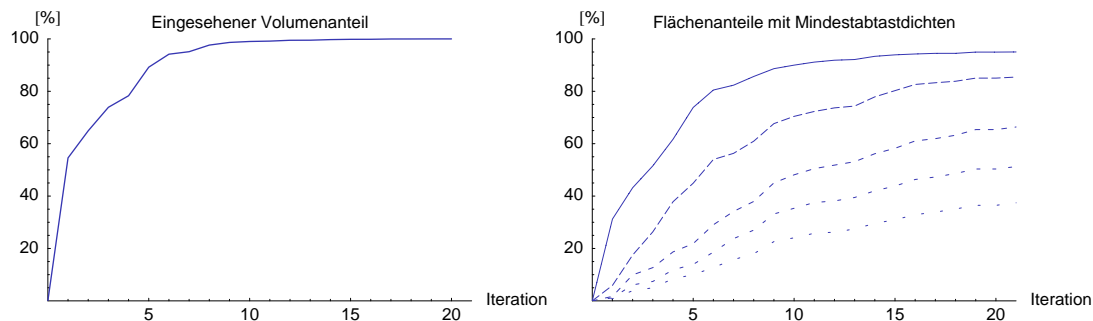


Abbildung 5.4: Ergebnisse für die künstliche Szene: Prozentsatz des eingesehenen Volumens (links) und Prozentsatz der Oberfläche, der mit einer Abtastdichte von ≥ 0 , $\geq 25\%$, $\geq 50\%$, $\geq 75\%$ und $\geq 100\%$ der gewünschten Abtastdichte repräsentiert ist (rechts, von oben nach unten).

Die simulierte Bildgeometrie und die Nebenbedingungen entsprechen denen des realen Akquisitionsgeräts (näherungsweise polare Bildgeometrie):

Höhe		0.9 m – 2.0 m
Blickrichtung	horizontal	$\pm 180^\circ$
	vertikal	$\pm 60^\circ$
Sensorgeometrie		Polarkoordinaten
Blickfeld	horizontal	$38^\circ - 278^\circ$
	vertikal	$30^\circ - 60^\circ$
Winkelauflösung		bis zu 0.1° / Pixel
gewünschte Abtastdichte		1600 Punkte / m ²

Dabei wird ein Abstand von wenigstens 60 cm zur nächsten Oberfläche vorausgesetzt, um das Akquisitionsgerät platzieren zu können.

Abbildung 5.4 zeigt die vorgeschlagenen Bewertungskriterien für eine Sequenz von 20 Aufnahmen. Während der Anteil des eingesehenen Volumens bereits in Iteration 12 99% des Gesamtvolumens erreicht, sind zum Erreichen der gewünschten Abtastdichte erheblich mehr Iterationen erforderlich: obwohl 95% der Oberfläche in einer der 20 Aufnahmen erfaßt werden geschieht dies nur für 37.4% mit der gewünschten Abtastdichte. Dies ist im Wesentlichen auf die komplizierte Szenengeometrie zurückzuführen, die eine große Zahl von Einzelaufnahmen aus kleinem Abstand zur Oberfläche erfordert. Das vorgeschlagene Verfahren zielt jedoch vorrangig auf die Auflösung noch nicht eingesehener Volumina ab. Dies wird in der Testszene auch tatsächlich erreicht; die verbliebenen 5% der Oberfläche, die in keiner der Aufnahmen repräsentiert sind, liegen im Wesentlichen auf der Unterseite der Tischplatten und ähnlichen Bereichen, die wegen der Einschränkungen der Akquisitionspositionen auf den Arbeitsbereich der Aufnahmeplattform unvermessbar sind.

5.3.1.2 Labor-Testszene

Die zweite Testszene ist eine typische Labor-Umgebung, die als Beispiel komplizierter, praxisrelevanter Szenen dient. Die Geometrie ist im Ergebnisbild in Abbildung 5.6 gut zu erkennen; sie enthält zwei verschiedene Deckenhöhen und eine verhältnismäßig schmale Nische. Etwa in der Mitte des Raums wurde die Situation durch einen stirnseitig am Arbeitstisch anschließenden Schrank künstlich verkompliziert. Neben diversen Stühlen, Tischen, Regalen und Rechnern (mit Anschlußkabeln) liefert ein von der Decke hängendes Trackingsystem aus 15 mm dicken Metallrohren interessante Details, die mit der gewünschten Abtastdichte nicht zuverlässig rekonstruiert werden können. Die technischen Parameter der Vermessung sind bis auf die gewünschte und die hardware-seitig mögliche Auflösung identisch zur im vorigen Abschnitt beschriebenen Situation:

Höhe		0.9 m – 2.0 m
Blickrichtung	horizontal	$\pm 180^\circ$
	vertikal	$\pm 60^\circ$
Sensorgeometrie		Polarkoordinaten
Blickfeld	horizontal	$38^\circ - 278^\circ$
	vertikal	$30^\circ - 60^\circ$
Winkelauflösung		bis zu 0.5° / Pixel
gewünschte Abtastdichte		625 Punkte / m ²

Abbildung 5.5 zeigt die vorgeschlagenen Evaluationskriterien der ersten 10 Aufnahmen. Während das rekonstruierte Volumen noch einige wichtige Verdeckungen ver-

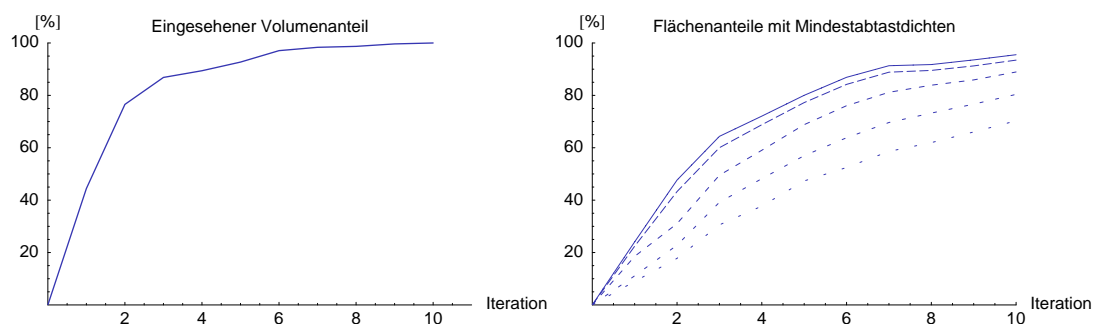


Abbildung 5.5: Ergebnisse für die reale Labor-Szene: Prozentsatz des eingesehenen Volumens (links) und Prozentsatz der Oberfläche, der mit einer Abtastdichte von ≥ 0 , $\geq 25\%$, $\geq 50\%$, $\geq 75\%$ und $\geq 100\%$ der gewünschten Abtastdichte repräsentiert ist (rechts, von oben nach unten). Alle Werte beziehen sich auf den finalen Zustand nach 10 Aufnahmen.

missen läßt, liefert das Verfahren auf mehr als 70% der rekonstruierten Oberfläche eine hinreichende Abtastdichte. Dies zeigt die Effizienz des Verfahrens im Fall einer limitierten Anzahl von Gesamtaufnahmen, die hier wegen der zur Verfügung stehenden Meßzeit die wesentliche Einschränkung bildete. Die Problematik einer nicht registrierbaren Einzelaufnahme trat weder während des hier beschriebenen Experiments noch während vorangegangener, hier nicht im Detail beschriebener Testläufe auf. Ebenso konnten die oben beschriebenen Details, die mit der gewünschten Abtastdichte nicht vermessbar sind, durch Bildfilterung erfolgreich entfernt werden, so daß sie den Planungsablauf nicht störten.

Abbildung 5.6 zeigt omnidirektional die sichtbare Oberfläche des eingesehenen Volumens für die geplante sechste Aufnahmeposition zusammen mit der Randlinie des geplanten Sichtfeldes und der nachfolgend tatsächlich akquirierten sechsten Aufnahme. In Abbildung 5.7 ist das rekonstruierte Ergebnis in mehreren computergraphisch generierten Ansichten zu sehen.

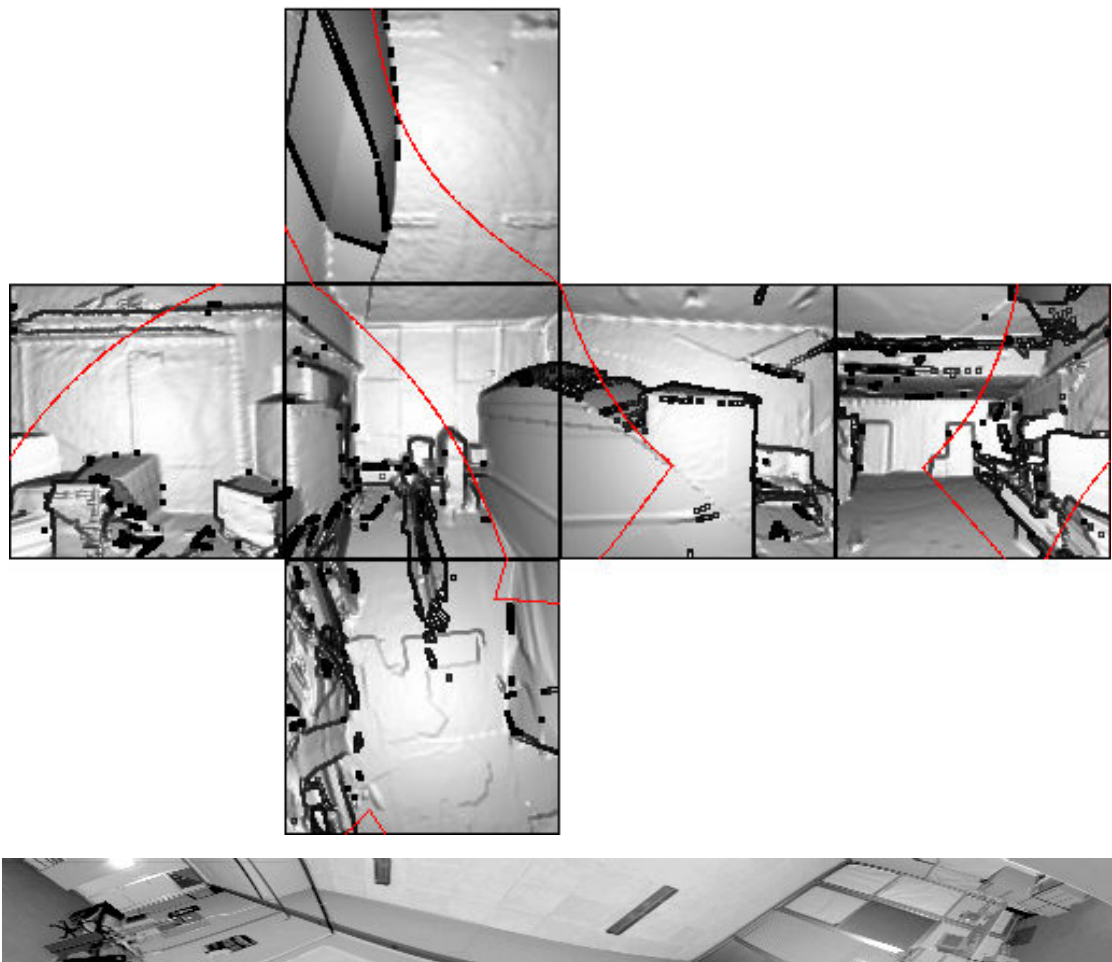


Abbildung 5.6: Gerenderte Darstellung der Oberfläche des eingesehenen Volumens auf den Flächen des view cubes mit der Randlinie des Sichtfeldes (rote Linie) (oben), sowie das Intensitätsbild der korrespondierenden Aufnahme (unten).

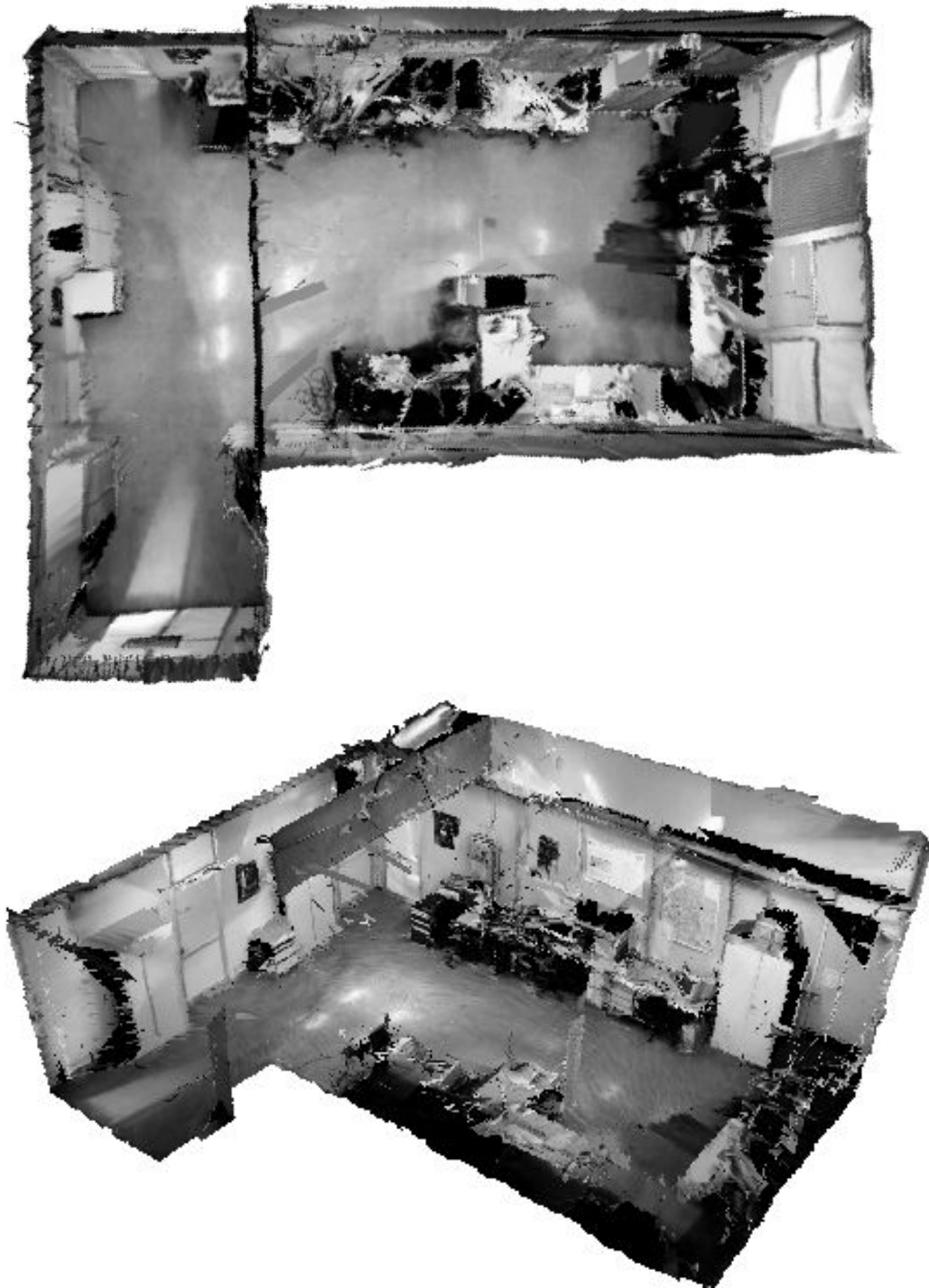


Abbildung 5.7: Zwei Ansichten des Rekonstruktionsergebnisses einer Labor-Umgebung nach 10 automatisch geplanten Aufnahmen. Die Grauwert-Textur wurde dabei aus den Intensitätsbildern gewonnen, die der verwendete Lichtlaufzeitsensor liefert.

5.3.2 Externe Rekonstruktion

Die praktische Anwendung des vorgeschlagenen Verfahrens für extern vermessene Objekte zeigt einige grundsätzliche Unterschiede zur internen Vermessung von Innenräumen. Während 3D-Meßsysteme, die auf Aufgaben der Gebäudevermessung spezialisiert sind, oftmals ein 360°-Blickfeld aufweisen und daher bereits mit der ersten Aufnahme eine weitgehende Übersicht über die vorliegende Geometrie der Szene genutzt werden kann, ist dies bei der externen Vermessung erst nach vielen Einzelaufnahmen möglich. Diesem Unterschied wird in der Praxis Rechnung getragen, indem zunächst eine ungeplante, automatisierte Sequenz von Aufnahmen mit konstanten Rotations-Schritten auf dem in das genutzte Roboter-System integrierten Drehteller akquiriert wird. Erst danach werden die verbliebenen Lücken der Oberflächen-Erfassung durch mit dem vorgeschlagenen Verfahren geplante Aufnahmen geschlossen und die gewünschte Abtastdichte sichergestellt.

Eine weiterer wichtiger Unterschied zwischen intern vermessenen Räumen und extern vermessenen Objekten ist die zunächst nicht einsehbare Standfläche. Während das mobile Akquisitionssystem zur Vermessung von Innenräumen sich innerhalb des vermessenen Objekts weitgehend frei bewegen kann, ist eine kontrollierte Lageänderung des auf dem Drehteller platzierten Objekts nicht möglich, sondern muß manuell ausgeführt werden. Dabei geht im Allgemeinen die bekannte relative Orientierung des

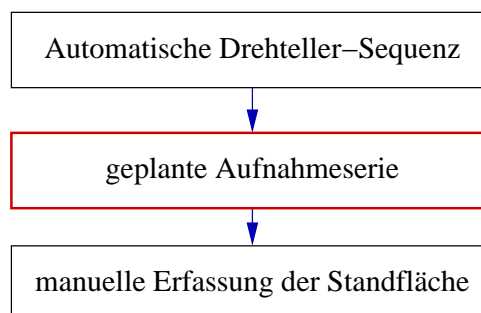


Abbildung 5.8: *Ablauf der automatisierten Vermessung von Objekten.*

Objekts zum Koordinatensystem des Meßsystems verloren, so daß für die nachfolgenden Aufnahmen keine Näherungslösung zur Integration in den Gesamtbildverbund zur Verfügung steht. Im nachfolgend beschriebenen Experiment wurden daher vier zusätzliche Aufnahmen manuell definiert und durch benutzerdefinierte Punktkorrespondenzen mit einer Näherungslösung der relativen Orientierung versehen, so daß abschließend auch Aufnahmen der Standfläche in den Bildverbund integriert werden konnten. Der Gesamtablauf der Vorgehensweise ist in Abbildung 5.8 dargestellt.

Das verwendete Beispielobjekt aus dem Automobil-Bereich repräsentiert einen typischen Fall der Einzelstückvermessung, der in der Praxis häufig auftritt. Es handelt sich

um ein bereits repariertes Originalteil eines historischen Rennwagens, für das keine Konstruktionsdaten mehr verfügbar sind. Diese Konstruktionsdaten sollen nun auf Basis der erfaßten Meßdaten durch *reverse engineering* gewonnen und zur Fertigung eines Ersatzteils verwendet werden.

Die Vermessung wurde mit den folgenden technischen Parametern durchgeführt:

Arbeitsbereich		Reis RV 16 Roboter mit Drehteller
Kamerageometrie		perspektivische Projektion
Triangulationswinkel		18.8°
Blickfeld	horizontal	22.6°
	vertikal	16.95°
Kameraauflösung		1024 × 768 Pixel
gewünschte Abtastdichte		15 Punkte / mm ²

Die gewünschte Abtastdichte trägt dabei dem Rekonstruktionsverfahren Rechnung, welches die redundante Information in mehrfach vermessenen Bereichen nutzt, um durch gewichtete Mittelwertbildung eine Erhöhung der Datenqualität und eine Eliminierung von Meßfehlern zu erreichen. Die native Auflösung der Kamera führt in der Mitte des kalibrierten Volumens zu einer Abtastdichte von nur bis zu 7 Punkten pro Quadratmillimeter.

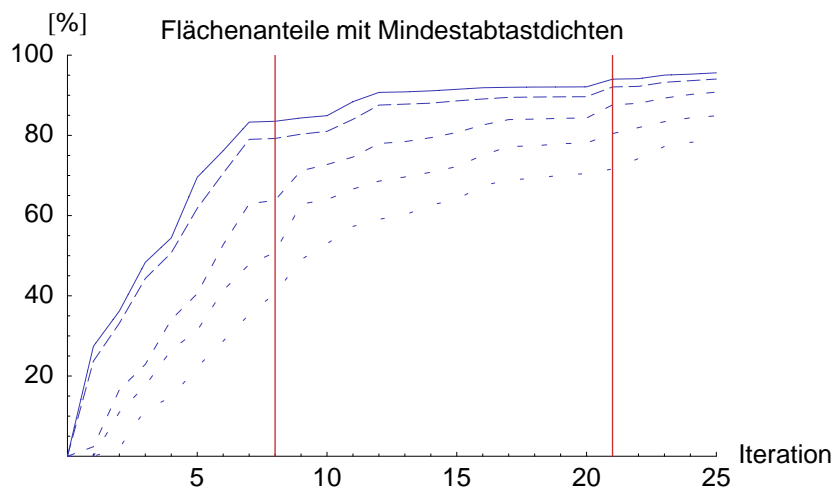


Abbildung 5.9: Ergebnisse für das Automobil-Werkstück: Prozentsatz der Oberfläche, der mit einer Abtastdichte von ≥ 0 , $\geq 25\%$, $\geq 50\%$, $\geq 75\%$ und $\geq 100\%$ der gewünschten Abtastdichte repräsentiert ist (von oben nach unten). Alle Werte beziehen sich auf den finalen Zustand nach 25 Aufnahmen, in dem verbliebene Lücken des Modells ergänzt wurden. Die roten Linien markieren den Übergang der in Abbildung 5.8 unterschiedenen Abschnitte der Vermessung.

Abbildung 5.9 zeigt das vorgeschlagene Evaluationskriterium „Abtastdichte“. Aus

den Daten ist zu erkennen, daß zunächst auch ohne gezielte Planung der Aufnahmen ein guter, 80% der Oberfläche umfassender Überblick über die Objektgeometrie gewonnen werden kann. Die Vermessung des verbleibenden Flächenanteils gestaltet sich zäh, aber der Anteil der Oberfläche, der mit der gewünschten Abtastdichte vermessen werden konnte, wird von 35% auf 75% gesteigert. Zu Beginn des geplanten Abschnitts tritt dabei das Problem auf, das zunächst noch nicht eingesehene Volumina als möglicherweise opak klassifiziert werden und so die Planung erschweren. In schlecht einsehbaren Bereichen können diese Volumina im Verlauf der Messung nicht aufgelöst werden, wie in Abbildung 5.10 zu sehen.

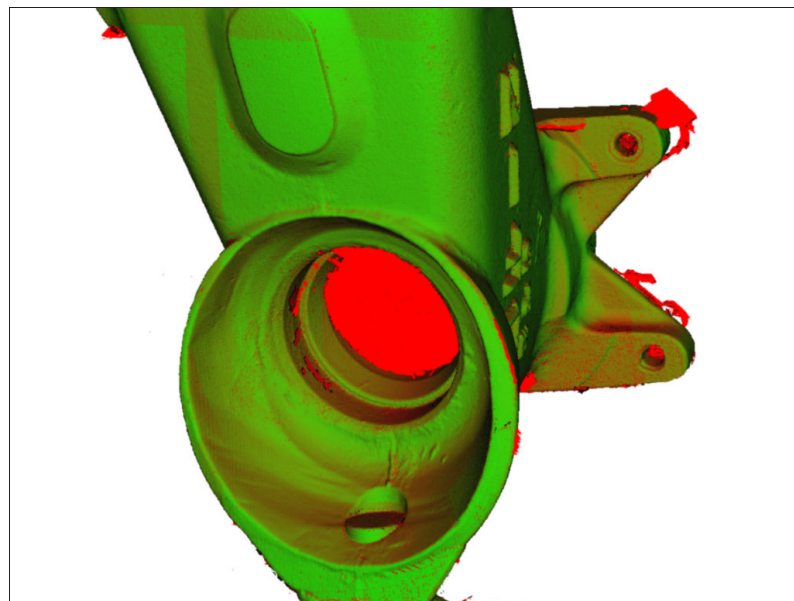


Abbildung 5.10: Ansicht der Objektrepräsentation am Ende der Planung. Der zylinderförmige Hohlraum konnte nicht als leer erkannt werden.

Zukünftige Arbeiten müssen an dieser Stelle zeigen, inwieweit mit nur einer Kamera entschieden werden kann, ob ein gegebenes Volumen in Bohrungen oder größeren zylinderförmigen Hohlräumen eingesehen wird. Als Ansatz kann dabei eine Durchlicht-Lösung dienen, wie sie in Abschnitt 3.5.2.2 skizziert wurde. In vielen Fällen typischer Industrie-Objekte wird jedoch nur die Einbeziehung von Benutzerwissen über die Objektgeometrie helfen, die interaktiv geschehen müßte.

Abbildung 5.12 zeigt die Sequenz der aufgenommenen Tiefenbilder. Die Oberfläche des Objekts konnte zu 95.6% erfaßt werden, dabei wird allerdings von einigen nicht einsehbaren Hohlräumen abgesehen. Zur Vereinfachung der weiteren Verarbeitung wird das Dreiecksnetz an unvermessenen Stellen initial durch die Grenzfläche des eingesehenen Volumens ergänzt; daraus werden lokal Minimalflächen zum Schließen der

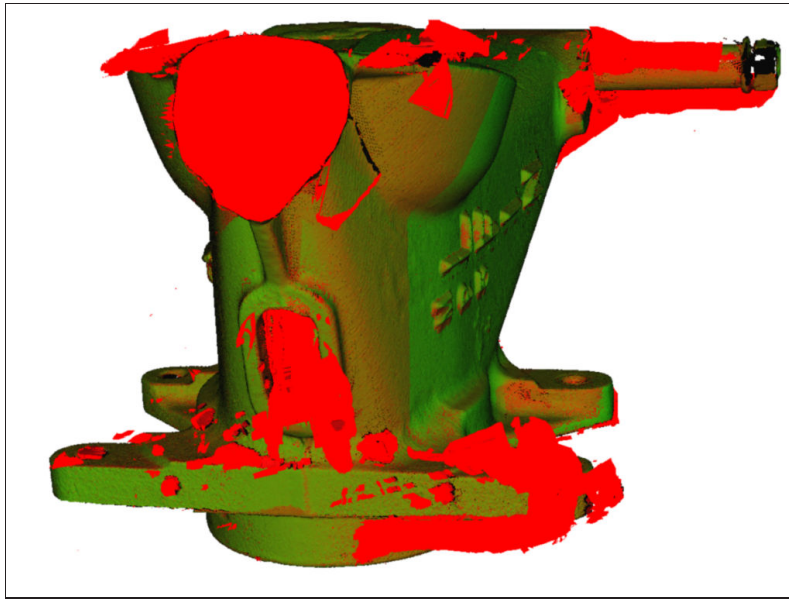


Abbildung 5.11: Planungsergebnis am Ende der Aufnahmesequenz nach Freigabe der durch den Arbeitsbereich gegebenen Nebenbedingungen: die Standfläche wird als nächste Ansicht vorgeschlagen.

Lücken abgeleitet, so daß im Ergebnis eine geschlossene, „wasserdichte“ Oberfläche entsteht. Auf deren Flächeninhalt beziehen sich alle Prozentangaben.

Die Abbildungen 5.13 und 5.14 zeigen das Ergebnis der Rekonstruktion. Die Ansichten sollen auch zeigen, daß die verwendete Geometrie für ungeübtes Bedienpersonal durchaus erhebliche Schwierigkeiten verursacht, wenn eine vorgegebene Ergebnisqualität mit geringem Aufwand erreicht werden soll.

5.4 Zusammenfassung eigener Ergebnisse

In diesem Kapitel wurden Maßzahlen vorgeschlagen, mit deren Hilfe die Ergebnisse einer Aufnahmesequenz quantifiziert und prinzipiell auch mit anderen Sequenzen, die nicht notwendigerweise mit demselben Verfahren geplant worden sein müssen, verglichen werden können. Hierzu wird Bild für Bild der Anteil des bereits eingesehenen Volumens sowie die Qualität der Oberflächenerfassung bezüglich der gewünschten Abtastdichte angegeben.

Weiterhin wurde das in der vorliegenden Arbeit vorgeschlagene Verfahren an drei Beispielen getestet und mit den in diesem Kapitel beschriebenen Maßzahlen evaluiert. Hierzu wurden eine künstliche und eine reale Innenraum-Szene sowie ein komplexes Objekt aus dem Automobil-Bereich verwendet.



Abbildung 5.12: Ansicht aller Scans des Beispielobjekts. Die Aufnahmen 0 bis 7 wurden mit einer ungeplanten Drehtellersequenz akquiriert, die letzten vier Aufnahmen (mit “downside” bezeichnet) wurden nach dem Umdrehen des Objekts halbmanuell positioniert.

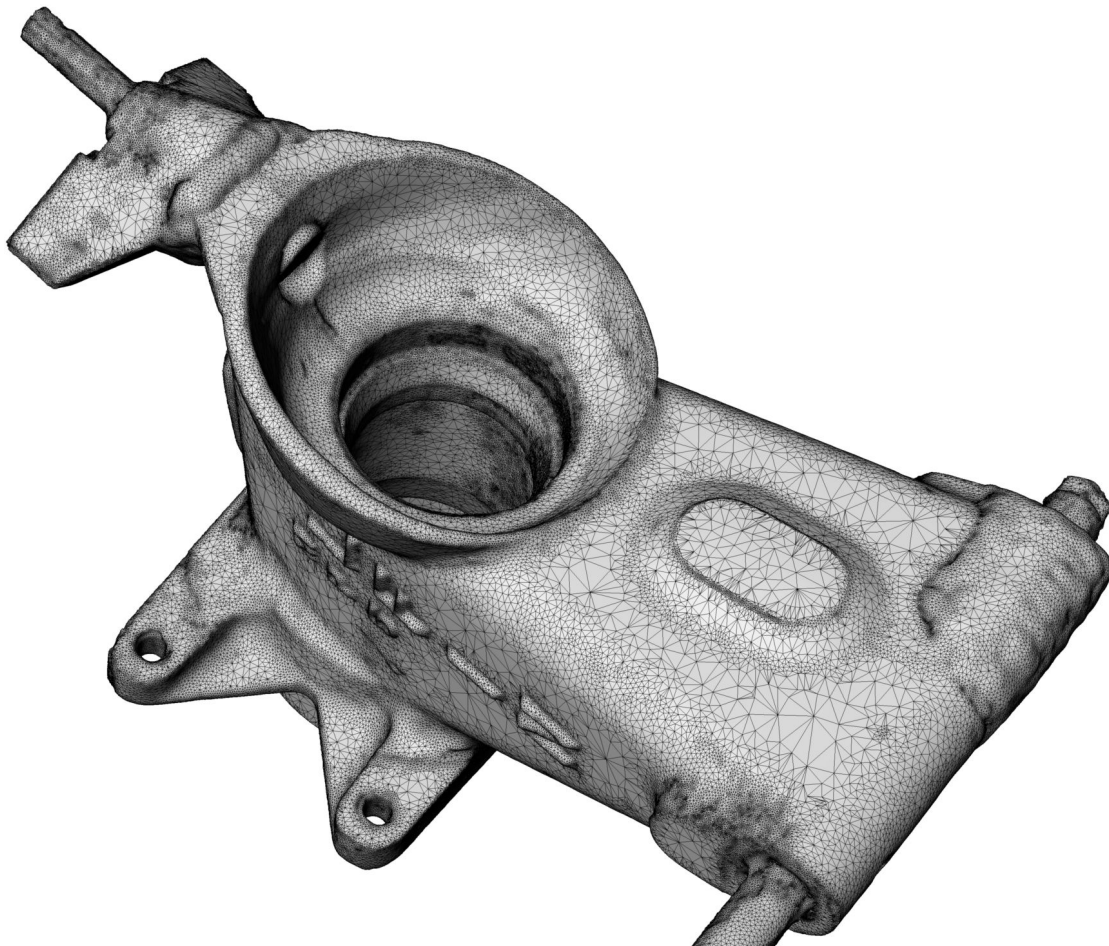


Abbildung 5.13: *Schattierte Darstellung der rekonstruierten Objektgeometrie mit überlagerten Dreieckskanten.*



Abbildung 5.14: Schattierte Darstellung des rekonstruierten Objekts mit farbig markierten Bereichen maximaler Krümmung.

KAPITEL 6

Zusammenfassung und Ausblick

Ziel der vorliegenden Arbeit war die Entwicklung eines Verfahrens zur automatisierten Rekonstruktion von Innenräumen und Gegenständen mit Hilfe eines 3D-Meßsystems und einem zu dessen Positionierung eingesetzten Roboter. Ein wesentlicher Anspruch bestand darin, diese Aufgabe auch für die in der Praxis oftmals auftretenden detailreichen und komplexen Geometrien zu lösen, die eine substantielle Anzahl von Einzelaufnahmen mit einer Vielzahl von frei wählbaren Aufnahmeparametern erfordern. Neben einer zielgerichteten Planung stand daher die Entwicklung effizient ablaufender Berechnungsmethoden im Vordergrund. Dazu war eine Reihe von Teilproblemen zu lösen: die Repräsentation der teilvermessenen Objektgeometrie, die Definition einer Zielfunktion zur vergleichenden Bewertung von Aufnahmekonfigurationen, sowie die Optimierung dieser Zielfunktion mit der Maßgabe, in jeder Iteration die jeweils beste nächste Aufnahmekonfiguration (*“next best view”*) zu finden.

Der Ansatz des vorgeschlagenen Verfahrens ist eine volumetrische Betrachtung des akquirierten Tiefenbildes, die neben den Meßdaten auf der Objektoberfläche auch die bei der Akquisition verwendeten Beobachtungspunkte auswertet. Diese Betrachtungsweise motiviert zunächst die Klassifikation der 3D-Meßverfahren nach der Anzahl der eingesetzten Beobachtungspunkte in unifokale, multifokale und triangulationsbasierte Verfahren. Durch Auswertung der Position der verwendeten Beobachtungspunkte kann der eingesehene und daher bekanntermaßen leere Teil des Raumes identifiziert werden. Die Grenzflächen dieses eingesehenen Volumens werden durch die vorgeschlagenen Verfahren modelliert, mit qualitätsbezogenen Attributen versehen und inkrementell ergänzt, wenn ein neu akquiriertes Bild in den Bildverbund integriert wird.

Auf dieser Repräsentation der teilvermessenen Objektgeometrie setzt die eigentliche Aufnahmeplanung auf, die auf einer Bewertungsfunktion basiert. Die Bewertungsfunktion berechnet für eine gegebene Aufnahmekonfiguration die Qualitätsverbesserung

rung der 3D-Rekonstruktion, die durch die entsprechende Aufnahme nach dem vorliegenden Kenntnisstand erreicht wird. Diese Aufgabe hat strukturell große Ähnlichkeit mit der computergraphischen Verarbeitungskette der 3D-Geometriedarstellung. Daher wurden die wesentlichen Berechnungsschritte systematisch auf computergraphische Basisfunktionen abgebildet, für die in der typischen Graphik-Hardware effiziente Implementierungen vorhanden sind.

Der wichtige Aspekt der effizienten Durchführung der Aufnahmeplanung kann mit dieser Strategie auf einfache Weise durch eine Anpassung der Rechengenauigkeit erreicht werden. So wurde von einer vereinfachten Geometrie des teilvermessenen Objekts ausgegangen, und es wurden Verfahren vorgeschlagen, die diese vereinfachte Geometrie bei der Integration zusätzlicher Aufnahmen erhalten. Dieser Ansatz wurde für die beiden wichtigsten computergraphischen Repräsentationsformen, nämlich Dreiecksnetze und Punktmengen, algorithmisch umgesetzt.

Weiterhin kann die Genauigkeit der Berechnung der Bewertungsfunktion im Rahmen der computergraphischen Berechnungsunterstützung durch die Wahl einer geeigneten Bildgröße gesteuert werden. Zusätzlich wurde für unifokale Meßsysteme gezeigt, wie die computergraphisch erzeugten Bilder, die die Grundlage der Funktionswertberechnung bilden, bei der aufeinanderfolgenden Evaluierung von Aufnahmekonfigurationen zum gleichen Beobachtungspunkt wiederverwertet werden können. Für triangulationsbasierte Sensoren wurde ein Verfahren zur computergraphischen Modellierung der Meßfeldgeometrie unter Verwendung von Verfahren der Schattenrechnung vorgeschlagen.

Eine weitere Steuerungsmöglichkeit ergibt sich durch die Wahl der Abtastdichte des Parameterraums bei der globalen Optimierung der Aufnahmekonfiguration mit Hilfe der Bewertungsfunktion. Weil die resultierende Laufzeit aber von den zu beachtenden Nebenbedingungen und nicht zuletzt auch von der Objektgeometrie abhängt, ist die Abtastdichte selbst nicht als Steuerungsparameter geeignet. Daher wurde eine Abtaststrategie vorgeschlagen, die laufzeitorientiert gesteuert werden kann, da zu jedem Zeitpunkt im Optimierungsablauf eine gleichmäßige Abtastung des Parameterraums sichergestellt wird.

Die vorgeschlagenen Verfahren wurden in zwei verschiedenen 3D-Rekonstruktionssystemen implementiert, getestet und evaluiert. Hierzu wurden im Rahmen der vorliegenden Arbeit geeignete Maßzahlen vorgeschlagen, die sensorunabhängig die Effizienz des Planungsverfahrens evaluieren und so prinzipiell auch einen Vergleich verschiedener Verfahren ermöglichen. Als Testfälle wurden eine künstliche und eine reale Innenraum-Szene sowie ein komplexes Objekt aus dem Automobil-Bereich ausgewählt.

Das vorgeschlagene Gesamtverfahren umfaßt einen vollständigen Satz von Lösungen für die oben genannten Teilprobleme, die mit dem Detailreichtum und der Kom-

plexität der zu rekonstruierenden Geometrie skalieren und effizient berechnet werden können. Im Unterschied zur verbreiteten Technik der Beschleunigung unveränderter Algorithmen durch eine sogenannte „GPU-Implementierung“ werden in der vorliegenden Arbeit strukturelle Ähnlichkeiten genutzt, um wesentliche Teile der Berechnung auf computergraphische Basisfunktionen zurückzuführen. Dabei werden nicht nur eine Verbesserung des Laufzeitverhaltens durch Ausführung der konkreten Berechnung im Graphik-Subsystem erreicht, sondern auch einfache Skalierungsmöglichkeiten durch Geometrievereinfachung und Reduktion der verwendeten Bildauflösung.

Möglichkeiten der zukünftigen Weiterentwicklung des vorgeschlagenen Verfahrens liegen im bisher wenig beachteten Zusammenspiel mit dem Positioniersystem. Abhängig von der erreichbaren Bewegungsgeschwindigkeit des Roboters kann eine Berücksichtigung des zurückzulegenden Weges bei der Bewertung von Aufnahmekonfigurationen ein erhebliches Beschleunigungspotential freisetzen (insbesondere für mobile Aufnahmesysteme).

Im Falle extern vermessener Gegenstände ist das Problem der Vermessung der Standfläche bisher ungelöst. Das Objekt muß manuell im Verlauf der Vermessung umgedreht werden, um auch die Unterseite erfassen zu können. Dieser Vorgang sollte in der Planung von Anfang an berücksichtigt werden, indem beispielsweise bestimmte Bereiche, die zunächst nur suboptimal vermessen werden können, zurückgestellt werden, um im zweiten Teil der Messung auf einfachere Weise erfaßt zu werden.

Ein erhebliches Vereinfachungspotential bei der Bestimmung der optimalen Aufnahmekonfiguration liegt in einer vorangehenden Analyse derjenigen Bereiche der Objektoberfläche, die durch die Integration der letzten Aufnahme unverändert bleiben. Beispielsweise wird sich die optimale Position für eine zusätzliche Aufnahme der Rückseite durch die Integration einer Aufnahme der Vorderseite im Regelfall nicht ändern; noch plakativer läßt sich die Situation bei der Innenraumrekonstruktion in einer Mehrraum-Umgebung vor Augen führen. Hier wäre eine Segmentierung in mehrere überlappende Teilbereiche mit Zwischenspeicherung der jeweils aussichtsreichsten lokalen Aufnahmekonfigurationen ein vielversprechender Ansatz.

Ein weiteres für die Praxis der Serien-Digitalisierung wichtiges Thema ist die Sicherheit von Objekt, Meßsystem und Roboter. Hier sind geeignete Ansätze zu entwickeln, die eine unabhängige und mit der geforderten Sicherheit fehlerfreie Prüfung einer errechneten Aufnahmekonfiguration in Bezug auf mögliche Kollisionen durchführen. Erst auf diese Weise wird die unbeaufsichtigte 3D-Digitalisierung unbekannter Objekte im Dienstleistungsalltag praktikabel.

ANHANG A

Formeln

A.1 Berechnung der Sensorauflösung aus der computergraphischen Ansicht

Hier soll der rechnerische Zusammenhang zwischen dem von einem Kamerapixel abgedeckten Raumwinkel und der entsprechenden Größe in einer computergraphischen Ansicht hergeleitet werden. Zur Berechnung des Raumwinkels der Kamera wird auf das in Abschnitt 3.4.4.1 beschriebene Kalkül auf Basis der Projektionsfunktion h und ihrer Jacobi-Matrix verwiesen. Analog zur Berechnung der Abtastdichte auf der vermessenen Oberfläche kann leicht hergeleitet werden, welcher Flächeninhalt dem abtastenden Pixel auf der Einheitskugel um den Beobachtungspunkt zuzuordnen ist.

Für die Bestimmung der entsprechenden Größe in einer computergraphischen Ansicht ist die inverse Projektion relevant, also die Abbildung einer Rasterposition (x_w, y_w) zusammen mit dem zugeordneten z -buffer-Eintrag z_w in die Koordinaten (x, y, z) der Szenenbeschreibung. Dieser Vorgang wird durch die 4×4 -Projektionsmatrix P und die 1×4 -viewport-Spezifikation (Fensterposition und -größe) v gesteuert und kann der Spezifikation zur OpenGL [98, 16] entnommen werden. Die Umformung erfolgt durch

$$\begin{pmatrix} x_c \\ y_c \\ z_c \\ w_c \end{pmatrix} = \begin{pmatrix} \frac{2(x_w - v_1)}{v_3} - 1 \\ \frac{2(y_w - v_2)}{v_4} - 1 \\ 2z_w - 1 \\ 1 \end{pmatrix} \quad (\text{A.1})$$

sowie die inverse Projektion mit

$$\begin{pmatrix} x_e \\ y_e \\ z_e \\ w_e \end{pmatrix} = P^{-1} \begin{pmatrix} x_c \\ y_c \\ z_c \\ w_c \end{pmatrix} \quad (\text{A.2})$$

und schließlich

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x_e/w_e \\ y_e/w_e \\ z_e/w_e \end{pmatrix} \quad (\text{A.3})$$

Unter Ausnutzung der Zwischenergebnisse ergibt sich die Jacobi-Matrix durch Differenzieren und Umformen zu

$$J = \begin{pmatrix} \frac{2(w_e p_{11} - x_e p_{41})}{w_e^2 v_3} & \frac{2(w_e p_{12} - x_e p_{42})}{w_e^2 v_4} & \frac{2(w_e p_{13} - x_e p_{43})}{w_e^2} \\ \frac{2(w_e p_{21} - y_e p_{41})}{w_e^2 v_3} & \frac{2(w_e p_{22} - y_e p_{42})}{w_e^2 v_4} & \frac{2(w_e p_{23} - y_e p_{43})}{w_e^2} \\ \frac{2(w_e p_{31} - z_e p_{41})}{w_e^2 v_3} & \frac{2(w_e p_{32} - z_e p_{42})}{w_e^2 v_4} & \frac{2(w_e p_{33} - z_e p_{43})}{w_e^2} \end{pmatrix} \quad (\text{A.4})$$

mit

$$P^{-1} =: \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \\ p_{41} & p_{42} & p_{44} & p_{44} \end{pmatrix} \quad (\text{A.5})$$

Damit sind die Voraussetzungen erfüllt, um jedem Pixel der graphischen Darstellung analog zur oben beschriebenen Vorgehensweise für die Kamera einen Raumwinkel zuzuordnen. Diese Raumwinkel können nun zueinander in Beziehung gesetzt werden, so daß im Ergebnis zu jedem Pixel der graphischen Darstellung die korrespondierende Auflösung der Kamera berücksichtigt werden kann.

Literaturverzeichnis

- [1] E.H. ADELSON and J.R. BERGEN: *The Plenoptic Function and the Elements of Early Vision*, chapter 1. MIT Press, Cambridge, Mass., 1991.
- [2] MARC ALEXA, JOHANNES BEHR, DANIEL COHEN-OR, SHACHAR FLEISHMAN, DAVID LEVIN, and CLAUDIO T. SILVA: *Computing and rendering point set surfaces*. IEEE Transactions on Visualization and Computer Graphics, 9(1):3–15, 2003.
- [3] K. ARIKAWA and S. HIROSE: *Study of walking robot for 3 dimensional terrain*. In *IEEE Int. Conf. on Robotics and Automation (ICRA '95)*, volume 1, pages 703 – 708, 1995.
- [4] A. ARSENIO and M. I. RIBEIRO: *Active range sensing for mobile robot localization*. In *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '98)*, volume 2, pages 1066 – 1071, 1998.
- [5] JAMES ARVO: *Fast random rotation matrices*. In D. KIRK (editor): *Graphics Gems III*, pages 117 – 120. Academic Press, 1992.
- [6] J. E. BANTA, Y. ZHIEN, X. Z. WANG, G. ZHANG, M. T. SMITH, and M. A. ABIDI: *A “best-next-view” algorithm for three-dimensional scene reconstruction using range images*. In *Intel. Robotics and Computer Vision XIV: Algorithms, Techniques, Active Vision, and Materials Handling*, pages 418–429, Philadelphia, 23–26 Oct. 1995. SPIE.
- [7] P. J. BESL: *Active Optical Range Imaging Sensors*. In J.L.C. SANZ (editor): *Advances in Machine Vision*, pages 1–63. Springer-Verlag, New York, 1989.
- [8] PAUL S. BLAER and PETER K. ALLEN: *View planning for automated site modeling*. In *IEEE Int. Conf. on Robotics and Automation (ICRA 2006)*, pages 2621–2626, Orlando, Florida, May 2006.
- [9] FRANÇOIS BLAIS: *Review of 20 years of range sensor development*. Journal of Electronic Imaging, 13(1):231–240, January 2004.

- [10] JULES BLOOMENTHAL and KEITH FERGUSON: *Polygonization of non-manifold implicit surfaces*. In *Computer Graphics - Annual Conference Series (SIGGRAPH '95 Proceedings)*, pages 309–316, 1995.
- [11] JOHN F. BROCK: *Pyramids to Pythagoras: Surveying from Egypt to Greece - 3000 B.C. to 100 A.D.* In *Proceedings FIG Working Week 2004*. International Federation of Surveyors (FIG), 2004.
- [12] ILJA NIKOLAEVIČ BRONSTEIN und KONSTANTIN ADOLFOVIČ SEMENDJAJEW: *Taschenbuch der Mathematik*. BSB Teubner, Leipzig, 24. Auflage, 1989.
- [13] DUANE C. BROWN: *Close-range camera calibration*. *Photogrammetric Engineering*, 37(8):855–866, 1971.
- [14] CARBON-VERTRIEB WEISSGERBER, WALLERSTEIN: *Technische Daten Verbundwerkstoff Kohlefaser mit Epoxidharzmatrix*. verfügbar unter <http://www.carbon-vertrieb.com>, September 2007.
- [15] FRANK CHEN, GORDON M. BROWN, and MUMIN SONG: *Overview of three-dimensional shape measurement using optical methods*. *Optical Engineering*, 39(1):10–22, 2000.
- [16] NORMAN CHIN, CHRIS FRAZIER, PAUL HO, ZICHENG LIU, and KEVIN P. SMITH: *The OpenGL graphics system utility library (version 1.3)*. Technical report, Silicon Graphics, Inc., 1998.
- [17] M. F. COHEN and D. P. GREENBERG: *The hemi-cube: A radiosity solution for complex environments*. In *Computer Graphics - Annual Conference Series (SIGGRAPH '85 Proceedings)*, pages 31–40, 1985.
- [18] C. I. CONOLLY: *The determination of next best views*. In *IEEE Int. Conf. on Robotics and Automation (ICRA '85)*, pages 432–435, 1985.
- [19] A. E. CONRADY: *Decentred lens-systems*. *Monthly Notices of the Royal Astronomical Society*, 79:384–390, March 1919.
- [20] JOHANNES G. VAN DER CORPUT: *Verteilungsfunktionen I + II*. In *Nederl. Akad. Wetensch. Proc.*, volume 38, pages 813–821, 1058–1066, 1935.
- [21] INGEMAR J. COX: *Blanche: Position estimation for an autonomous robot vehicle*. In *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '89)*, pages 432 – 439, 1989.
- [22] MICHAEL CRAMER, DIRK STALLMANN, and NORBERT HAALA: *Sensor integration and calibration of digital airborne three-line camera systems*. In *Proc. Int. Workshop on Mobile Mapping Technology*, pages 4–5–1 – 4–5–8, Bangkok, Thailand, April 1999.

- [23] FRANKLIN C. CROW: *Shadow algorithms for computer graphics*. In *Computer Graphics - Annual Conference Series (SIGGRAPH '77 Proceedings)*, pages 242–248, 1977.
- [24] BRIAN CURLESS and MARC LEVOY: *A volumetric method for building complex models from range images*. In *Computer Graphics - Annual Conference Series (SIGGRAPH '96 Proceedings)*, pages 303–312, 1996.
- [25] CYBERWARE, INC.: *Rapid 3d digitizing platform scans the human head*. Press release, Monterey, CA, November 22 1991. available at <http://www.cyberware.com> (April 2008).
- [26] J. DENAVIT and R. S. HARTENBERG: *A kinematic notation for lower-pair mechanisms based on matrices*. ASME Journal of Applied Mechanics, pages 215–222, 1955.
- [27] THOMAS D. DEWITT and DOUGLAS A. LYON: *A range finding method using diffraction gratings*. Applied Optics, 34(14):2510–2521, May 1995.
- [28] S. DUBOWSKY and E. PAPADOPOULOS: *The kinematics, dynamics, and control of free-flying and free-floating space robotic systems*. IEEE Transactions on Robotics and Automation, 9(5):531 – 543, 1993.
- [29] FRÉDO DURAND: *A multidisciplinary survey of visibility*. In *ACM Siggraph course notes "Visibility, Problems, Techniques, and Applications"*, 2000.
- [30] CASS EVERITT, ASHU REGE, and CEM CEBENOYAN: *Hardware shadow mapping*. White paper, available at http://developer.nvidia.com/object/hwshadowmap_paper.html (April 2008), NVidia Corporation, Santa Clara, CA, USA, 2001.
- [31] JAMES D. FOLEY, ANDRIES VAN DAM, STEVEN K. FEINER, and JOHN F. HUGHES: *Computer Graphics – Principles and Practice*. Addison Wesley, Reading, Massachusetts, 2nd edition, 1992.
- [32] J. G. FRYER: *Camera calibration*. In ATKINSON (editor): *Close Range Photogrammetry and Machine Vision*, pages 156–179. Whittles Publishing, Caithness, UK, 1996.
- [33] NATASHA GELFAND, LESLIE IKEMOTO, SZYMON RUSINKIEWICZ, and MARC LEVOY: *Geometrically stable sampling for the ICP algorithm*. In *Proceedings of Fourth International Conference on 3-D Digital Imaging and Modeling (3DIM'03)*, page 260, 2003.
- [34] GESELLSCHAFT FÜR OPTISCHE MESSTECHNIK MBH, BRAUNSCHWEIG: *Dokumentation zum Meßsystem ATOS III* verfügbar unter <http://www.gom.com>, April 2008.

- [35] NED GREENE: *Environment mapping and other applications of world projections*. IEEE Computer Graphics and Applications, 6(11):21–29, November 1986.
- [36] MICHAEL D. GROSSBERG and SHREE K. NAYAR: *A general imaging model and a method for finding its parameters*. In *Proc. International Conference on Computer Vision (ICCV'01)*, volume 2, pages 108–115, 2001.
- [37] MICHAEL D. GROSSBERG and SHREE K. NAYAR: *The raxel imaging model and ray-based calibration*. International Journal of Computer Vision, 61(2):119–137, 2005.
- [38] V. I. GUSHOV and YU. N. SOLODKIN: *Automatic processing of fringe patterns in integer interferometers*. Optics and Lasers in Engineering, 14:311–324, 1991.
- [39] OLAF HALL-HOLT: *Scrutinizing real-world geometry: the next best view*. Technical report, Stanford University, Stanford, CA, USA, 1998.
- [40] J. H. HALTON: *On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals*. Numer. Math., 2:84–90, 1960.
- [41] J. M. HAMMERSLEY: *Monte-carlo methods for solving multivariable problems*. Ann. New York Acad. Sci., 86:844–874, 1960.
- [42] RICHARD HARTLEY and SING BING KANG: *Parameter-free radial distortion correction with center of distortion estimation*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 29(8):1309 – 1321, 2007.
- [43] RICHARD HARTLEY and ANDREW ZISSERMAN: *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2001.
- [44] JEAN-MARC HASENFRATZ, MARC LAPIERRE, NICOLAS HOLZSCHUCH, and FRANÇOIS SILLION: *A survey of real-time soft shadows algorithms*. In *Computer Graphics Forum (EUROGRAPHICS 2003 Proceedings)*. Eurographics, 2003. State-of-the-Art Report.
- [45] HERMANN HEISS: *Grundlagen der Koordinatentransformation bei Industrierobotern*. In: *Robotersysteme 2*, Seiten 65–71. Springer, 1986.
- [46] H. HÖFLER, C. BAULIG, A. BLUG, M. DAMBACHER, N. DIMOPOULOS, and H. WÖLFELSCHNEIDER: *Optical high-speed 3d metrology in harsh environments: Recording structural data of railway lines*. In *Proc. Conf. Optical Measurement Systems for Industrial Inspection*, SPIE Proceedings Series 5856, pages 296–306, Munich, June 13-16 2005.
- [47] TAKEO KANADE, ATSUSHI YOSHIDA, KAZUO ODA, HIROSHI KANO, and MASAYA TANAKA: *A stereo machine for video-rate dense depth mapping and its new applications*. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR '96)*, page 196, San Francisco, 1996.

- [48] MICHAEL KAZHDAN, MATTHEW BOLITHO, and HUGUES HOPPE: *Poisson surface reconstruction*. In *Eurographics Symposium on Geometry Processing*, 2006.
- [49] G. S. KIRK and J. E. RAVEN: *The Presocratic Philosophers*. Cambridge University Press, 1957.
- [50] H. KITANO, M. FUJITA, S. ZREHEN, and K. KAGEYAMA: *Sony legged robot for robocup challenge*. In *IEEE Int. Conf. on Robotics and Automation (ICRA '98)*, volume 3, pages 2605 – 2612, Leuven, May 16–20 1998.
- [51] KONRAD KLEIN and VÍTOR SEQUEIRA: *The view-cube: An efficient method of view planning for 3d modelling from range data*. In *Proceedings 5th IEEE Workshop on Applications of Computer Vision (WACV 2000)*, pages 186–191, 2000.
- [52] KONRAD KLEIN and VÍTOR SEQUEIRA: *View planning for the 3d modelling of real world scenes*. In *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2000)*, pages 943–948, 2000.
- [53] KONRAD KLEIN and VÍTOR SEQUEIRA: *View planning for unknown indoor scenes based on a cost benefit analysis*. In 23. *Symposium für Mustererkennung - Jahrestagung der Deutschen Arbeitsgemeinschaft für Mustererkennung e.V. (DAGM 2001)*, number 2191 in *Springer Lecture Notes in Computer Science (LNCS)*, pages 313–320, Munich, 2001.
- [54] JAMES J. KUFFNER: *Effective sampling and distance metrics for 3D rigid body path planning*. In *IEEE Int. Conf. on Robotics and Automation (ICRA 2004)*, pages 3993 – 3998, New Orleans, LA, 2004.
- [55] JEAN-CLAUDE LATOMBE: *Robot Motion Planning*. Kluwer Academic Publishers, Boston, Dordrecht, London, 1991.
- [56] ALDO LAURENTINI: *Inferring the shape of the real object from the object reconstructed by volume intersection*. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR '93)*, pages 280–285, New York, 1993.
- [57] STEVEN M. LAVALLE: *Planning Algorithms*. Cambridge University Press, 2006.
- [58] ERIC LENGYEL: *The OpenGL Extensions Guide*. Charles River Media, 2003.
- [59] M. LHUILLIER and LONG QUAN: *Robust dense matching using local and global geometric constraints*. In *Proceedings 15th International Conference on Pattern Recognition*, volume 1, pages 968–972, Barcelona, September 2000.
- [60] F. LORANGER, D. LAURENDEAU, and R. HOUDE: *A fast and accurate 3-D range finder using the Biris technology: the TRID sensor*. In *Proc. Intern. Conf. on Recent Advances in 3-D Digital Imaging and Modeling*, pages 51 – 58, 1997.
- [61] W.E. LORENSEN and H. E. CLINE: *Marching cubes: A high resolution 3d surface construction algorithm*. *Computer & Graphics*, 21(4), 1987.

- [62] THOMAS LUHMANN: *Nahbereichsphotogrammetrie: Grundlagen, Methoden und Anwendungen*. Wichmann, Heidelberg, 2000.
- [63] W. N. MARTIN and J. K. AGGARWAL: *Volumetric descriptions of objects from multiple views*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 5(2):150 – 158, 1983.
- [64] NIKOLAOS A. MASSIOS and ROBERT B. FISHER: *A best next view selection algorithm incorporating a quality criterion*. In *British Machine Vision Conference*, pages 780–789, 1998.
- [65] JASNA MAVER and RUZENA BAJCSY: *Occlusions as a guide for planning the next view*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 15(5):417–433, May 1993.
- [66] MAYTEC ALUMINIUM SYSTEMTECHNIK GMBH, DACHAU: *Technische Daten zu Strangpreßprofilen nach DIN 17615*. verfügbar unter http://www.maytec.de/technische_daten.htm, April 2008.
- [67] LEONARD MCMILLAN and GARY BISHOP: *Plenoptic modeling: An image-based rendering system*. In *Computer Graphics - Annual Conference Series (SIGGRAPH '95 Proceedings)*, 1995.
- [68] M. MINSKY: *Memoir on inventing the confocal scanning microscope*. Scanning, 10:128–138, 1988.
- [69] C. MONTANI, R. SCATENI, and R. SCOPIGNO: *A modified look-up table for implicit disambiguation of marching cubes*. Visual Computer, 10(6), 1994.
- [70] SHREE K. NAYAR, MASAHIRO WATANABE, and MINORI NOGUCHI: *Real-time focus range sensor*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 18(12):1186–1198, December 1996.
- [71] PETER J. NEUGEBAUER: *Reconstruction of real-world objects via simultaneous registration and robust combination of multiple range images*. International Journal of Shape Modeling, 3(1 & 2):71–90, 1997.
- [72] PETER J. NEUGEBAUER: *3D-Digitalisierung und Rekonstruktion realer Objekte unter Berücksichtigung der Sensorgeometrie*. Dissertation, Technische Universität Darmstadt, 2000.
- [73] PETER J. NEUGEBAUER and KONRAD KLEIN: *Adaptive triangulation of objects reconstructed from multiple range images*. In *IEEE Visualization '97, Late Breaking Hot Topics*, Phoenix, Arizona, 1997.
- [74] PETER J. NEUGEBAUER und KONRAD KLEIN: *Adaptive Triangulierung komplexer Oberflächen*. In: 5. ABW-Workshop, TA Esslingen, Januar 1999.

- [75] PETER J. NEUGEBAUER and KONRAD KLEIN: *Texturing 3d models of real world objects from multiple unregistered photographic views*. In *Computer Graphics Forum (EUROGRAPHICS '99 Proceedings)*, volume 18, pages 245 – 256, 1999.
- [76] H. NIEDERREITER: *Random Number Generation and Quasi-Monte-Carlo Methods*. Society for Industrial and Applied Mathematics, Philadelphia, USA, 1992.
- [77] K. L. PAAP, T. CHRISTALLER, and F. KIRCHNER: *A robot snake to inspect broken buildings*. In *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2000)*, volume 3, pages 2079 – 2082, 2000.
- [78] D. PIEPER and B. ROTH: *The kinematics of manipulators under computer control*. In *Proc. 2nd int. Congr. on Theory of Machines and Mechanisms*, pages 159–169, Zakopane, Poland, 1969.
- [79] RICHARD PITO: *A sensor based solution to the next best view problem*. In *Int. Conference on Pattern Recognition*, pages 941–945, Vienna, Austria, August 1996.
- [80] RICHARD PITO: *A solution to the next best view problem for automated surface acquisition*. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(10):1016 – 1030, 1999.
- [81] PMD TECHNOLOGIES GMBH, SIEGEN. verfügbar unter <http://www.pmdtec.com>, April 2008.
- [82] MARC POLLEFEYS, REINHARD KOCH, MAARTEN VERGAUWEN, and LUC VAN GOOL: *Metric 3d surface reconstruction from uncalibrated image sequences*. In R. KOCH and L. VAN GOOL (editors): *3D Structure from Multiple Images of Large-Scale Environments: European Workshop, SMILE'98*, volume 1506 of *Lecture Notes in Computer Science*. Springer, June 1998.
- [83] KARI PULLI: *Multiview registration for large data sets*. In *Proceedings of Second International Conference on 3D Digital Imaging and Modeling (3DIM'99)*, pages 160–168, Ottawa, Canada, October 1999.
- [84] CHRISTIAN RATHJEN: *Mehr-Wellenlängen-Verfahren unter dem Aspekt von Parametervariationen und der Separation von Parametern*. In: *3. ABW Workshop, TA Esslingen*, Esslingen, 1996.
- [85] CHRISTIAN RATHJEN: *Untersuchungen zu Mehr-Wellenlängen-Verfahren und deren Anwendungen auf Projektionsmoirés zur absoluten Formerrfassung*. Fortschritt-Berichte VDI: Reihe 8, Meß-, Steuerungs- und Regelungstechnik; 579 (Dissertation, Universität Hannover). VDI-Verlag, 1996.
- [86] MICHAEL K. REED and PETER K. ALLEN: *Constraint-based sensor planning for scene modeling*. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12):1460–1467, December 2000.

- [87] MICHAEL K. REED, PETER K. ALLEN, and IOANNIS STAMOS: *3-d modeling from range imagery: An incremental method with a planning component*. In *International Conference on Recent Advances in 3-D Digital Imaging and Modeling*, pages 76–83, Ottawa, Canada, 1997.
- [88] REIS GMBH & CO. KG MASCHINENFABRIK, OBERNBURG: *Dokumentation zum Vertikalknickarmroboter RV16*. verfügbar unter <http://www.reisrobotics.de>, April 2008.
- [89] D. R. ROBERTS and A. D. MARSHALL: *Viewpoint selection for complete surface coverage of three dimensional objects*. In *British Machine Vision Conference*, pages 740–750, 1998.
- [90] F. ROCKER: *Localisation and classification of 3d objects*. In *Proceedings 2nd Intern. Conf. on Pattern Recognition*, pages 527–528, Copenhagen, Denmark, August 1974.
- [91] JÁNOS ROHÁLY and DOUGLAS P. HART: *High resolution, ultra fast 3-D imaging*. In BRIAN D. CORNER and JOSEPH H. NURRE (editors): *Three-Dimensional Image Capture and Application III*, volume 3958 of *Proceedings of SPIE*, pages 2–10, 2000.
- [92] K. SATO and S. INOKUCHI: *3D surface measurement by space encoding range imaging*. *Journal of Robotic Systems*, 2(1):27–39, 1985.
- [93] D. SCHARSTEIN and R. SZELISKI: *A taxonomy and evaluation of dense two-frame stereo correspondence algorithms*. *International Journal of Computer Vision*, 47(1/2/3):7–42, April - June 2002.
- [94] H. SCHEWE and W. FÖRSTNER: *The program PALM for automatic line and surface measurement using image matching techniques*. In *Symposium Intern. Soc. for Photogrammetry and Remote Sensing*, volume 26, pages 608–622, 1986.
- [95] WOLFGANG SCHREIBER and GUNTHER NOTNI: *Theory and arrangements of self-calibrating whole-body three-dimensional measurement systems using fringe projection technique*. *Optical Engineering*, 39(1):159–169, 2000.
- [96] WILLIAM R. SCOTT, GERHARD ROTH, and JEAN-FRANÇOIS RIVEST: *View planning as a set covering problem*. Technical Report NRC-44892, National Research Council of Canada, Institute for Information Technology, Ottawa, Ont., Canada, 2001.
- [97] WILLIAM R. SCOTT, GERHARD ROTH, and JEAN-FRANÇOIS RIVEST: *View planning for automated three-dimensional object reconstruction and inspection*. *ACM Computing Surveys*, Vol. 35(1):64–96, March 2003.
- [98] MARK SEGAL and KURT AKELEY: *The OpenGL graphics system: A specification (version 1.5)*. Technical report, Silicon Graphics Inc., 2003.

- [99] VÍTOR SEQUEIRA, JOÃO G. M. GONÇALVES, and M. ISABEL RIBEIRO: *Active view selection for efficient 3d scene reconstruction*. In *Proc. ICPR'96 - 13th International Conference on Pattern Recognition*, pages 815–819, Vienna, Austria, 1996.
- [100] B. W. SILVERMAN: *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, London, 1986.
- [101] CHESTER C. SLAMA, CHARLES THEURER, and SOREN W. HENRIKSEN (editors): *Manual of Photogrammetry*. American Soc. of Photogrammetry and Remote Sensing, Falls Church, Virginia, 4th edition, 1980.
- [102] GIDEON P. STEIN: *Lens distortion calibration using point correspondences*. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR '97)*, pages 602–608, 1997.
- [103] PETER F. STURM and STEPHEN J. MAYBANK: *On plane-based camera calibration: A general algorithm, singularities, applications*. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR '99)*, 1999.
- [104] HORST STÖCKER (Herausgeber): *Taschenbuch der Physik: Formeln, Tabellen, Übersichten*. Verlag Harri Deutsch, Thun, Frankfurt am Main, 2. völlig überarb. und erw. Auflage, 1994.
- [105] KONSTANTINOS A. TARABANIS, PETER K. ALLEN, and ROGER Y. TSAI: *A survey of sensor planning in computer vision*. IEEE Transactions on Robotics and Automation, 11(1):86–104, February 1995.
- [106] G. H. TARBOX and S. N. GOTTSCHLICH: *Planning for complete sensor coverage in inspection*. Computer Vision and Image Understanding, 61(1):84–111, January 1995.
- [107] SETH TELLER: *Automatic acquisition of hierarchical, textured 3d geometric models of urban environments: Project plan*. In *Proc. Image Understanding Workshop*, 1997.
- [108] HANS J. TIZIANI and HANS MARTIN UHDE: *Three dimensional image sensing with chromatic confocal microscopy*. Applied Optics, 33:1838–1843, 1994.
- [109] ROGER Y. TSAI: *A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses*. IEEE Transactions on Robotics and Automation, RA-3(4):323–344, 1987.
- [110] G. TURK and M. LEVOY: *Zippered polygon meshes from range images*. In *Computer Graphics - Annual Conference Series (SIGGRAPH '94 Proceedings)*, pages 311–318, Orlando, Florida, 1994.
- [111] MICHAEL VORMWALD: *Baumechanik II*. Skript zur Vorlesung im SS 2002, Bauhaus Universität Weimar, 2002.

- [112] WOLFGANG WEBER: *Industrieroboter – Methoden der Steuerung und Regelung*. Carl Hanser Verlag, München Wien, 2002.
- [113] J. WENG, P. COHEN, and M. HERNIOU: *Camera calibration with distortion models and accuracy evaluation*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 14(10):965–980, 1992.
- [114] PETER WHAITE and FRANK P. FERRIE: *From uncertainty to visual exploration*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 13(10):1038–1049, October 1991.
- [115] ANDREW WOO, PIERRE POULIN, and ALAIN FOURNIER: *A survey of shadow algorithms*. IEEE Computer Graphics and Applications, 10(6):13–32, 1990.
- [116] Z. XU, R. SCHWARTE, H. HEINOL, B. BUXBAUM, and T. RINGBECK: *Smart pixel – photonic mixer device (pmd)*. In *Proc. Int. Conf. on Mechatron. and Machine Vision*, 1998.
- [117] ANNA YERSHOVA and STEVEN M. LAVALLE: *Deterministic sampling methods for spheres and SO(3)*. In *IEEE Int. Conf. on Robotics and Automation (ICRA 2004)*, 2004.
- [118] HANSONG ZHANG: *Forward shadow mapping*. In *Rendering Techniques 1998 (9th Eurographics Workshop on Rendering)*, pages 131–138, 1998.
- [119] HANSONG ZHANG: *Forward shadow mapping*. Technical Report TR98-003, University of North Carolina at Chapel Hill, Department of Computer Science, 1998.

Lebenslauf

Angaben zur Person

KONRAD FRIEDRICH KLEIN
geboren am 17. Februar 1972 in Würzburg
deutsche Staatsangehörigkeit



Ausbildung und beruflicher Werdegang

- 1978 – 1982 Grundschule
- 1982 – 1991 Gymnasium
- 1991 – 1998 Studium der Informatik an der Technischen Universität Darmstadt; Abschluß als Diplom-Informatiker.
- 1998 – 1999 Marie-Curie-Stipendiat an der Gemeinsamen Forschungsstelle der Europäischen Kommission in Ispra, Italien.
- 1999 – 2006 Wissenschaftlicher Mitarbeiter am Fraunhofer Institut für Graphische Datenverarbeitung, Darmstadt.
- 2004 – 2006 stellvertretender Abteilungsleiter der Abteilung "Cognitive Computing & Medical Imaging" am Fraunhofer IGD.
- seit 2006 Tätigkeit in der Forschung und Entwicklung bei Sirona Dental Systems GmbH, Bensheim.
- 2008 Promotion an der Technischen Universität Darmstadt.